



# تبدیل صفحک: یک موجک هندسی مناسب برای فشرده‌سازی تصاویر عمق دوربین‌های شبه کینکت

وحید کیانی، احد هراتی\* و عابدین واحدیان  
دانشکده مهندسی، دانشگاه فردوسی مشهد، مشهد، ایران

## چکیده

تبدیلات گوک (Wedgelet) و شبیک (Platelet) که پیش‌ازاین در خانواده موجک‌های هندسی و فقی برای بازنمایی تصاویر روشنایی مطرح شده‌اند، توانایی بازنمایی تنک تصاویر قطعه‌ای ثابت، و تصاویر قطعه‌ای خطی را دارند؛ اما کارایی آن‌ها در بازنمایی تصاویر قطعه‌ای غیرخطی مانند تصاویر عمق بهینه نیست. در این مقاله تبدیل صفحک<sup>۱</sup> به‌عنوان عضو جدیدی از خانواده موجک‌های هندسی برای بازنمایی بهینه تصاویر عمق قطعه‌ای صفحه‌گون ارائه شده است. برخلاف موجک‌های هندسی پیشین که تنها از مدل‌های خطی و ثابت برای توصیف هر ناحیه هموار در تصویر استفاده می‌کنند، تبدیل صفحک برای تقریب هر ناحیه صفحه‌گون از یک مدل غیرخطی مبتنی بر توابع گویا بهره می‌گیرد. آزمایش‌ها بر روی تصاویر عمق واقعی نشان دادند که در نرخ بیت ۰/۰۳ bpp استفاده از کدگذار مبتنی بر صفحک در فشرده‌سازی تصاویر عمق نسبت به موجک هندسی گوک به‌طور میانگین تا ۲/۷ dB کیفیت را افزایش می‌دهد. همچنین در شرایط مشابه، در مقایسه با کدگذارهای مدرن JPEG2000 و H.264، استفاده از کدگذار عمق مبتنی بر صفحک به‌ترتیب منجر به ۲/۵۹ dB و ۱/۵۶ dB افزایش در کیفیت تصاویر بازسازی شده می‌شود.

واژگان کلیدی: فشرده‌سازی تصویر عمق، موجک هندسی، تبدیل صفحک، توابع گویا.

## Planelet Transform: A New Geometrical Wavelet for Compression of Kinect-like Depth Images

Vahid Kiani, Ahad Harati\* & Abedin Vahedian

Computer Engineering Department, Ferdowsi University of Mashhad, Mashhad, Iran

### Abstract

With the advent of cheap indoor RGB-D sensors, proper representation of piecewise planar depth images is crucial toward an effective compression method. Although there exist geometrical wavelets for optimal representation of piecewise constant and piecewise linear images (i.e. wedgelets and platelets), an adaptation to piecewise linear fractional functions which correspond to depth variation over planar regions is still missing. Such planar regions constitute major portions of the indoor depth images and need to be well represented to allow for desirable rate-distortion trade-off.

In this paper, second-order planelet transform is introduced as an optimal representation for piecewise planar depth images with sharp edges along smooth curves. Also, to speed up the computation of planelet approximation of depth images, an iterative estimation procedure is described based on non-linear least squares and discontinuity relaxation. The computed approximation is fed to a rate-distortion optimized quad-tree based encoder; and the pruned quadtree is encoded into the bit-stream. Spatial horizontal and vertical

<sup>1</sup> Planelet Transform

\* Corresponding author

\*نویسنده عهده‌دار مکاتبات

plane prediction modes are also introduced to further exploit geometric redundancy of depth images and increase the compression ratio.

Performance of the proposed planelet-based coder is compared with wedgelets, platelets, and general image encoders on synthetic and real-world Kinect-like depth images. The synthetic images dataset consists of 30 depth images of different scenes which are manually selected from eight video sequences of ICL-NUIM RGBD Benchmark dataset. The dataset of real-world images also includes 30 depth images of indoor scenes selected from Washington RGBD Scenes V2 dataset captured by Kinect-like cameras.

In contrast to former geometrical wavelets which approximate smooth regions of each image using constant and linear functions, planelet transform exploits a non-linear model based on linear fractional functions to approximate every smooth region. Visual comparisons by 3D surface reconstruction and visualization of the decoded depth images as surface plots revealed that at a specific bit-rate the planelets-based coder better preserves the geometric structure of the scene compared with the former geometric wavelets and the general images coders.

Numerical evaluations showed that compression of synthetic depth-images by planelets results in a considerable PSNR improvement of 0.83 dB and 6.92 dB over platelets and wedgelets, respectively. Due to absence of the noise, the plane prediction modes were very successful on synthetic images and boosted the PSNR gap over platelets and wedgelets to 5.73 dB and 11.82 dB, respectively. The proposed compression scheme also performed well on the real-world depth images. Compared with wedgelets, planelets-based coder with spatial prediction achieved noticeable quality improvement of 2.7 dB at the bit-rate of 0.03 bpp. It also led to 1.46 dB quality improvement over platelets at the same bit-rate. In this experiment, application of planelets-based coder led to 2.59 dB and 1.56 dB increase in PSNR over JPEG2000 and H.264 general image coders. Similar results are also achieved in terms of SSIM metric.

**Keywords:** depth-image compression, geometrical wavelets, planelet transform, linear fractional functions.

حفظ مرزهای اشیا و حفظ سطوح آن‌ها مورد بررسی قرار گیرد [9-11]. به‌منظور بازسازی دقیق مرزهای اشیا لازم است

لبه‌های تیز تصویر عمق تا حد ممکن حفظ شوند؛ از طرف دیگر، بازسازی بهتر سطوح اشیا نیازمند حفظ محتوای هر ناحیه هموار در تصویر عمق است. بنابراین، روش‌هایی که به حفظ لبه‌ها و حفظ نواحی هموار در تصاویر عمق توجه داشته باشند، در حفظ هندسه صحنه حین فشرده‌سازی نیز موفق خواهند بود.

کدگذارهای مدرن تصویر که در حال حاضر برای فشرده‌سازی عموم تصاویر به کار می‌روند، در حفظ لبه‌های تیز تصاویر عمق، عملکرد مناسبی نداشته و منجر به تولید نقاط غیرواقعی<sup>10</sup> زیادی در اطراف لبه‌ها می‌شوند [9, 10, 12, 13]. همچنین، باوجوداین که داده‌های عمق اخذشده از دوربین‌های عمق شبه‌کینکت، به‌طورعمومی حاوی نوفه نیز هستند، کدگذارهای مدرن تصاویر طبیعی نسبت به نوفه موجود در داده‌های عمق آگاهی نداشته و در کنار حفظ اطلاعات عمق بر حفظ نوفه نیز تأکید دارند. به همین دلیل، به روش‌های بهتری برای فشرده‌سازی تصاویر عمق نیاز است که ضمن تلاش برای حفظ هندسه، در مقابل نوفه نیز مقاوم باشند [8].

پیش‌ازاین، خانواده موجک‌های هندسی<sup>11</sup> توسط پژوهش‌گران به‌منظور حفظ بهتر لبه‌های تصویر در کاربردهای

## ۱- مقدمه

اخذ اطلاعات عمق از صحنه در حل مسائل ادراک صحنه<sup>1</sup>، هدایت ربات<sup>2</sup>، اجتناب از موانع<sup>3</sup>، بازشناسی اشیا<sup>4</sup>، تولید نقشه محیط و مکان‌یابی ربات<sup>5</sup>، و بازسازی سه‌بعدی صحنه<sup>6</sup> کاربرد دارد [1-5]. فناوری ارزان‌قیمت نور ساخت‌یافته<sup>7</sup> که در سال‌های اخیر عرضه شده، در حال حاضر رایج‌ترین رویکرد برای اخذ تصاویر عمق در محیط‌های داخلی<sup>8</sup> است [6, 7]. برای ذخیره‌سازی حجم انبوه تصاویر اخذشده از دوربین عمق، یا ارسال این تصاویر از طریق کانال‌های انتقال بی‌سیم با پهنای باند محدود به روش‌هایی برای فشرده‌سازی تصاویر عمق نیاز است [8]. برخلاف تصاویر رنگی که هدف اصلی در فشرده‌سازی آن‌ها دستیابی به کیفیت ادراکی بالاتر از دیدگاه انسان است، در فشرده‌سازی تصاویر عمق هدف اصلی دستیابی به کیفیت هندسی<sup>9</sup> بالاتر و حفظ اجزای اصلی صحنه است. نمونه‌ای از یک تصویر عمق و تصویر رنگ متناظر با آن در شکل (۱) نشان داده شده است.

حفظ هندسه یک صحنه سه‌بعدی می‌تواند از دو جنبه

<sup>1</sup> scene understanding

<sup>2</sup> robot guidance

<sup>3</sup> collision avoidance

<sup>4</sup> object recognition

<sup>5</sup> SLAM

<sup>6</sup> scene reconstruction

<sup>7</sup> structured-light technology

<sup>8</sup> indoor scenes

<sup>9</sup> geometric quality

<sup>10</sup> veil points

<sup>11</sup> geometrical wavelets

در این تبدیل، هر ناحیه از تصویر عمق به‌کمک یک مدل غیرخطی بازنمایی خواهد شد که برای توصیف مقادیر عمق حاصل از افکنش<sup>9</sup> یک صفحه مناسب است.

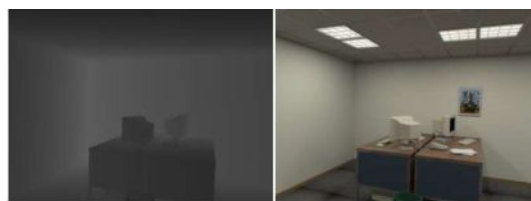
در دومین بخش این مقاله، مروری بر کارهای شاخص در زمینه فشرده‌سازی اطلاعات عمق ارائه می‌شود. در بخش سوم، مروری بر تبدیل گوک خواهیم داشت؛ سپس، در بخش چهارم تبدیل صفحه‌ک برای بازنمایی تَنک تصاویر عمق معرفی شده، و نحوه محاسبه سریع آن در بخش پنجم بیان می‌شود. در بخش ششم این مقاله، نحوه کدگذاری تصاویر عمق به‌کمک تبدیل صفحه‌ک تشریح خواهد شد. در بخش هفتم، روش‌های پیش‌بینی افقی و عمودی بلاک‌های عمق برای دستیابی به کارایی بالاتر در کدگذار عمق مبتنی بر صفحه‌ک معرفی خواهند شد؛ سپس، در بخش هشتم، عملکرد روش فشرده‌سازی پیشنهادی در فشرده‌سازی تصاویر عمق به‌کمک آزمایش‌های عملی ارزیابی شده و با موجک‌های هندسی پیشین و برخی کدگذارهای مدرن فشرده‌سازی تصاویر عمومی مقایسه می‌شود؛ در نهایت، در بخش نهم، مهم‌ترین یافته‌های این پژوهش در فشرده‌سازی تصاویر عمق به‌کمک موجک‌های هندسی جمع‌بندی شده و راه‌کارهایی برای توسعه این پژوهش در آینده پیشنهاد داده خواهد شد.

## ۲- پیشینه پژوهش

پیش از ارائه دوربین‌های عمق مبتنی بر نور ساخت‌یافته، دوربین‌های عمق دودیدی<sup>10</sup> که از دو دوربین رنگی برای محاسبه مقادیر عمق استفاده می‌کنند، رایج‌ترین ابزار برای اخذ تصاویر عمق بودند [21, 22]. در این دوربین‌ها برای محاسبه مقدار عمق هر پیکسل از تصویر، ناهنجاری<sup>11</sup> آن پیکسل بین تصویر دید راست و تصویر دید چپ محاسبه شده و تصویری از مقادیر ناهنجاری<sup>12</sup> تولید می‌شد. هنوز در برخی کاربردها که به تصاویر عمقی با جزئیات بالا و مقادیر عمق دقیق نیاز است، و مواردی که با صحنه‌های خارجی در فضای باز مواجه هستیم، دوربین‌های عمق دودیدی مورد استفاده قرار می‌گیرند.

تصاویر ناهنجاری محاسبه‌شده در دوربین‌های دودیدی در نواحی صفحه‌گون، طبیعت خطی دارند [9, 23]. از طرف دیگر تبدیل شیبک در خانواده موجک‌های هندسی برای بازنمایی بهینه تصاویر قطعه‌ای خطی با مرزهای تیز ارائه شده

فشرده‌سازی و تقریب تصاویر طبیعی معرفی شده‌اند [14-17]. موجک‌های هندسی وقتی<sup>1</sup> برای دستیابی به یک بازنمایی تَنک از تصویر، لبه‌های تصویر را به‌طور صریح و به‌کمک اتم‌های یک لغت‌نامه چنددقتی<sup>2</sup> توصیف می‌کنند؛ به‌علاوه، این تبدیل‌ها برای توصیف هر ناحیه در تصویر از مدل مناسبی بر اساس نوع تصویر استفاده می‌کنند. در این خانواده از تبدیل‌ات، تبدیل گوک<sup>3</sup> برای بازنمایی تصاویر قطعه‌ای یکنواخت<sup>4</sup> [14]، و تبدیل شیبک<sup>5</sup> برای بازنمایی تصاویر قطعه‌ای خطی<sup>6</sup> [18] رفتاری بهینه داشته و از کارایی بالایی برخوردار هستند. با این وجود، هنگامی که با تصاویر عمق سر و کار داریم، به‌دلیل طبیعت غیرخطی مقادیر عمق، عملکرد تبدیلات گوک و شیبک زیر بهینه خواهد بود.



(شکل-۱): نمونه‌ای از یک تصویر عمق و تصویر رنگ متناظر با آن در یک صحنه داخلی

(Figure-1): Sample depth image and its corresponding RGB image captured from an indoor scene

از آنجایی که دوربین‌های عمق مبتنی بر نور ساخت‌یافته تنها می‌توانند برای اخذ تصاویر عمق در محیط‌های داخلی مورد استفاده قرار گیرند، بخش اعظم پیکسل‌های تصاویر عمق اخذشده توسط آن‌ها شامل بازنمایی سطوح صفحه‌گون صحنه خواهد بود [1, 19, 20]. هنگامی که یک سطح صفحه‌گون بر روی تصویر عمق افکنده می‌شود، مقادیر عمق پیکسل‌های ناحیه حاصل با موقعیت آن پیکسل‌ها رابطه‌ای غیرخطی خواهند داشت؛ به همین دلیل تبدیلات گوک و شیبک که قبلاً در خانواده موجک‌های هندسی مطرح شده‌اند، از بازنمایی کارآمد سطوح صفحه‌گون در تصاویر عمق عاجز مانده، و در بازنمایی تصاویر عمق قطعه‌ای صفحه‌گون<sup>7</sup> عملکرد زیربهینه خواهند داشت. بر همین اساس، در این مقاله تبدیل صفحه‌ک<sup>8</sup> در خانواده موجک‌های هندسی، به‌عنوان توسعه مناسبی بر تبدیل گوک برای بازنمایی کارآمد تصاویر عمق معرفی شده و عملکرد آن در فشرده‌سازی تصاویر عمق ارزیابی خواهد شد.

<sup>1</sup> adaptive geometrical wavelets

<sup>2</sup> multi-resolution dictionary

<sup>3</sup> wedgelet transform

<sup>4</sup> piecewise constant images

<sup>5</sup> platelet transform

<sup>6</sup> piecewise linear images

<sup>7</sup> piecewise planar images

<sup>8</sup> planclct transform

<sup>9</sup> projection

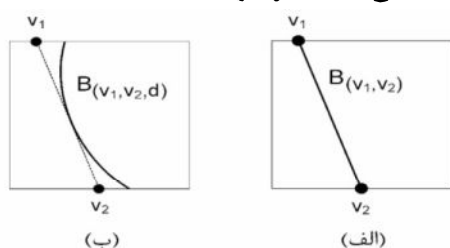
<sup>10</sup> stereo depth cameras

<sup>11</sup> disparity

<sup>12</sup> disparity image

[31]. هنگامی که تصاویر عمق متوالی از یک صحنه مشترک گرفته می‌شوند، دارای افزونگی زمانی بالایی هستند. بر همین اساس، در [8] از مجتمع‌سازی حجمی<sup>۳</sup> نقاط صحنه در یک شبکه سه‌بعدی برای تولید تقریبی از تصاویر عمق متوالی استفاده شده و سپس، از کدگذار H.264 برای فشرده‌سازی مقادیر باقی‌مانده پیش‌بینی بهره گرفته شده است. هر دو رویکرد تجمیع حجمی و پیش‌بینی یک‌بعدی به تولید یک تقریب خوب از مقادیر عمق منجر می‌شوند؛ اما این روش‌ها، راه‌کاری برای کدگذاری مستقل<sup>۴</sup> مقادیر عمق یک ناحیه ارائه نمی‌دهند؛ در مقابل، به‌منظور کدگذاری مستقل مقادیر عمق هر ناحیه و با هدف حفظ هندسه در فشرده‌سازی تصاویر عمق، در [32] برای تقریب هر بلاک تصویر عمق از یک صفحه در فضای سه‌بعدی استفاده شده است. از آنجایی که در این روش راه‌کاری برای مدل‌سازی دقیق لبه‌های تیز<sup>۵</sup> تصاویر عمق مطرح نشده، عملکرد آن نسبت به کدگذار H.264 پایین‌تر است.

در این مقاله، موجک هندسی صفحک به‌عنوان عضو جدیدی از خانواده موجک‌های هندسی برای فشرده‌سازی تصاویر عمق اخذشده از محیط‌های داخلی مطرح خواهد شد. در این تبدیل به‌منظور بازنمایی هر ناحیه در تصویر عمق از یک مدل غیرخطی<sup>۶</sup> مبتنی بر توابع گویا<sup>۷</sup> استفاده خواهد شد که برای بازنمایی نواحی صفحه‌گون مناسب است؛ به‌علاوه بازنمایی صفحک برای هر بلاک از تصویر عمق، علاوه‌بر دو ناحیه صفحه‌گون، حاوی مشخصات دقیق مرز بین این دو ناحیه نیز هست. بر این اساس، انتظار می‌رود که روش پیشنهادی در فشرده‌سازی تصاویر عمق، با حفظ بهتر لبه‌های تیز و نواحی هموار به حفظ بهتر هندسه صحنه و کیفیت بالاتر تصاویر کدگشایی‌شده منجر شود.



(شکل-۲): بازنمایی لبه به‌کمک انواع مختلف مرزک در موجک‌های هندسی و فقی: (الف) مرزک درجه یک، (ب) مرزک درجه دو

(Figure-2): Representation of block's edge in adaptive geometrical wavelets by (a) a first-order beamlet, and (b) a second-order beamlet

<sup>3</sup> volumetric integration

<sup>4</sup> intra-coding

<sup>5</sup> step/sharp edges

<sup>6</sup> non-linear model

<sup>7</sup> linear fractional functions

است [18]؛ به همین دلیل، نخستین تلاش‌ها برای استفاده از موجک‌های هندسی در فشرده‌سازی تصاویر عمق، به فشرده‌سازی تصاویر نابه‌جایی حاصل از دوربین‌های عمق دودیدی به‌کمک موجک هندسی شیبک باز می‌شود [9, 24]. این رویکرد در فشرده‌سازی ویدئوی چندمنظری MVD<sup>۱</sup> به‌کار گرفته شده، و به‌دلیل حفظ بهتر هندسه در تصاویر عمق دیده‌های اصلی، در بازسازی مبتنی بر عمق تصاویر رنگی دیده‌های میانی<sup>۲</sup> عملکرد بهتری را نسبت به H.264 ارائه داده است. استفاده از تبدیل شیبک در فشرده‌سازی تصاویر نابه‌جایی با توسعه تبدیل شیبک به مرزهای درجه دو نیز ادامه یافته است [25]. در کنار موجک‌های هندسی، گروهی از پژوهش‌گران برای فشرده‌سازی تصاویر نابه‌جایی، سعی کرده‌اند تا نسخه‌های اصلاح‌شده‌ای از روش‌های سنتی فشرده‌سازی تصویر مطرح کنند که نسبت به نسخه سنتی عملکرد بهتری را در حفظ لبه‌های تیز و نواحی خطی داشته باشند. ازجمله این رویکردها می‌توان به کاهش طول فیلترهای موجک سنتی در نزدیکی لبه [12, 26]، پس‌پردازش تصاویر عمق کدگشایی‌شده برای ترمیم لبه‌های عمق [27, 28]، و افزاز هر بلاک به‌کمک یک منحنی به دو ناحیه [29] اشاره کرد.

دوربین‌های عمق مبتنی بر بینایی دودیدی به زمان قابل توجهی برای محاسبه مقادیر نابه‌جایی بر اساس تصاویر رنگی اخذشده نیاز دارند [22]. انجام این پردازش توسط همان رایانه‌ای که دوربین به آن متصل است، باعث کندشدن عملکرد کلی سامانه شده و در بسیاری از کاربردهای بلادرنگ در حوزه رباتیک به یک چالش بزرگ تبدیل می‌شود. در مقابل، دوربین‌های عمق ارزان‌قیمت مبتنی بر نور ساخت‌یافته محاسبات عمق را در پردازنده داخلی خود انجام داده و امکان محاسبه و اخذ بلادرنگ تصاویر عمق را بدون متحمل شدن بار پردازشی اضافه فراهم می‌آورند [30]. به همین دلیل، در بسیاری کاربردها که با محیط‌های داخلی سر و کار داریم، دوربین‌های عمق مبتنی بر نور ساخت‌یافته به‌تدریج جایگزین دوربین‌های عمق مبتنی بر بینایی دو دیدی شده‌اند.

پژوهش‌های انجام‌شده برای فشرده‌سازی تصاویر عمق اخذشده از دوربین‌های نور ساخت‌یافته در سال‌های اخیر رشد قابل توجهی داشته‌اند. در یک رویکرد ساده برای فشرده‌سازی مقادیر عمق، با توجه به همواربودن نواحی صفحه‌گون در تصاویر عمق از روش پیش‌بینی ساده بین پیکسل‌های هم‌جوار برای فشرده‌سازی بدون تلف تصاویر عمق استفاده شده است

<sup>1</sup> multiple view plus depth (MVD)

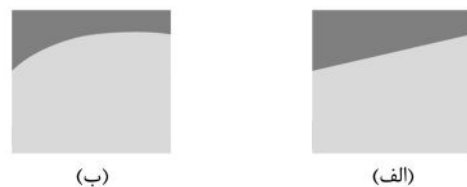
<sup>2</sup> intermediate views

## ۳- مروری بر تبدیل گوک

شالوده اصلی موجک‌های هندسی وفقی نحوه بازنمایی لبه توسط آن‌ها در هر بخش از تصویر است. در ابتدایی‌ترین رویکرد، موجک‌های هندسی وفقی مکان پیکسل‌های یک لبه تیز در بلاکی از تصویر را با کمک یک پاره‌خط صاف بازنمایی می‌کنند؛ که نقاط پایانی آن بین پیکسل‌های مرزی بلاک قرار دارد. به چنین لبه‌ای در اصطلاح «مرزک درجه یک»<sup>۱</sup> گفته می‌شود [33, 34]. نمونه‌ای از مرزک درجه یک در شکل (۲) (الف) نشان داده شده است. با کمک مفهوم مرزک درجه یک، بازنمایی تنک لبه در یک بلاک از تصویر با کدگذاری نقاط شروع و پایان مرزک امکان‌پذیر می‌شود. استفاده از روش‌های مختلف برای بازنمایی مقادیر پیکسل‌ها در هر یک از دو ناحیه حاصل و مرز بین آن‌ها، و پیچیده‌تر کردن بازنمایی لبه باعث به‌وجود آمدن موجک‌های هندسی وفقی مختلفی شده است.

## ۳-۱- تبدیل گوک درجه یک

تبدیل گوک درجه یک در سال ۱۹۹۸ توسط دیوید دونوهو به‌عنوان یک تبدیل چنددقتی برای بازنمایی تصاویر مطرح شد [14]. برای ایجاد یک بازنمایی چنددقتی از تصویر، تبدیل گوک از یک درخت چهارگانه استفاده می‌کند؛ تا تصویر را مرتب به بخش‌های کوچک‌تری تقسیم کند؛ سپس، هر بلاک از تصویر به کمک توابع پایه لغت‌نامه چنددقتی گوک تقریب زده می‌شود. در این بازنمایی، هر گره از درخت چهارگانه بلاک متناظرش در تصویر را تنها با یک عنصر از لغت‌نامه گوک توصیف می‌کند. هر عنصر پایه گوک، شامل یک مرزک است که بلاک را به دو ناحیه هم‌جوار تقسیم می‌کند. در تبدیل گوک، برای هر یک از این دو ناحیه، مقادیر تمام پیکسل‌های ناحیه به کمک یک مقدار روشنایی توصیف شده و با یک تابع مقدار ثابت تقریب زده می‌شوند [35]. نمونه‌ای از تقریب یک بلاک توسط یک عنصر گوک در شکل (۳) (الف) نمایش داده شده است.



(شکل-۳): تقریب یک بلاک به کمک: (الف) گوک درجه یک، (ب)

گوک درجه دو

(Figure-3): Approximation of a sample block by (a) a first-order wedgelet, and (b) a second-order wedgelet

<sup>۱</sup> first-order beamlet

تقریب گوک درجه یک برای یک بلاک از تصویر می‌تواند با رابطه زیر بیان شود:

$$M_{W(v_1, v_2, h_1, h_2)}(x, y) = h_1 + (h_2 - h_1)W_{(v_1, v_2)}(x, y) \quad (1)$$

که در آن پارامترهای  $(v_1, v_2)$  اندیس نقاط پایانی خط صاف لبه را نشان می‌دهند و به‌عنوان اندیس یک عنصر از لغت‌نامه گوک عمل می‌کنند؛ درحالی‌که، پارامترهای  $(h_1, h_2)$  مقادیر روشنایی پیکسل‌های هر یک از دو ناحیه حاصل را نشان می‌دهند، و عنصر گوک را تزئین می‌کنند. سمبل  $W_{(v_1, v_2)}$  ناحیه‌ای از بلاک را نشان می‌دهد که توسط گوکی با نقاط پایانی  $(v_1, v_2)$  مشخص شده است.

در یک بلاک  $M \times M$  از تصویر، نقاط مرزی بلاک می‌توانند با اندیس‌های ۰ تا  $4M - 1$  شماره‌گذاری شوند. در این صورت، مرزک درجه یک می‌تواند با کمک اندیس نقطه شروع و اندیس نقطه پایان آن به صورت  $b(v_1, v_2)$  مشخص شود؛ که در آن  $v_1 < v_2$  است. مجموعه مرزک‌هایی که به این شکل در یک بلاک تعریف می‌شوند، تمام لبه‌های ممکن در آن بلاک را پوشش نمی‌دهند؛ اما، وجود لبه‌هایی در مکان‌ها، مقیاس‌ها، و جهت‌های مختلف در لغت‌نامه چند دقتی مرزک موجب می‌شود تا این لغت‌نامه برای تصاویری با لبه‌های هموار  $C^\alpha$  توانایی بازنمایی لبه‌های تصویر را با دقت بالایی داشته باشد [14].

لغت‌نامه گوک برای تقریب یک بلاک شامل تمام گوک‌های ممکن برای بلاک‌هایی با آن اندازه است، که در آن تمام گوک‌های تباهیده<sup>۲</sup> با هم ادغام شده و به‌عنوان یک عنصر خاص از لغت‌نامه نمایش داده می‌شوند. گوک‌های تباهیده توانایی تقسیم بلاک به دو ناحیه را نداشته، و کل پیکسل‌های بلاک را به‌صورت یک ناحیه بازنمایی می‌کنند<sup>۳</sup>. بنابراین، لغت‌نامه گوک برای تمامی بلاک‌های  $M \times M$  تصویر که در سطح یکسانی از درخت چهارگانه قرار دارند، به شکل زیر تعریف می‌شود [35]:

$$W(M) = \{1_{M \times M}\} \cup \{\text{all } W_{(v_1, v_2)} \text{ of size } M \times M\} \quad (2)$$

لغت‌نامه گوک برای بازنمایی چنددقتی تصویر از ادغام لغت‌نامه‌های مربوط به سطوح مختلف درخت چهارگانه حاصل می‌شود و می‌تواند به شکل زیر تعریف شود:

<sup>۲</sup> degenerate beamlet

<sup>۳</sup> برای بعضی ترکیبات  $(v_1, v_2)$  از نقاط مرزی، مرزک درجه یک بر روی یکی از ضلع‌های بلاک قرار می‌گیرد. در این صورت، مرزک مذکور توانایی تقسیم بلاک به دو ناحیه مجزا را نداشته و مرزک تباهیده نامیده می‌شود.

$$W = \{W_{(j,i,m)}: j = 0, \dots, \log_2 N; \quad (3)$$

$$i = 0, \dots, 4^i - 1; \\ m = 0, \dots, \left\lfloor W\left(\frac{N}{2^i}\right) \right\rfloor - 1 \}$$

که در آن سمبل  $j$  نشان‌دهنده اندیس سطح در درخت چهارگانه، سمبل  $i$  نشان‌دهنده اندیس بلاک در آن سطح، و سمبل  $m$  نشان‌دهنده اندیس تابع پایه گوک در لغت‌نامه آن سطح است.

### ۲-۳- تبدیل گوک درجه‌دو

تبدیل گوک درجه دو در سال ۲۰۰۷ توسط آگنیسکا لیزوفسکا به‌عنوان نسخه‌ای از تبدیل گوک برای بازنمایی بهتر لبه‌های منحنی در کنار لبه‌های صاف پیشنهاد داده شد [34]. در تبدیل گوک درجه‌دو، به‌منظور بازنمایی بهتر لبه‌های منحنی، به‌جای یک پاره‌خط صاف از یک چندجمله‌ای درجه دو برای مدل‌سازی مکان پیکسل‌های لبه استفاده شده و با در نظر گرفتن میزان انحنای آن نسبت به یک پاره‌خط صاف، مفهوم «مرکز درجه دو» مطرح شده است. نمونه‌ای از یک مرکز درجه دو در شکل (۲) (ب) نشان داده شده است. نمونه‌ای از تقریب یک بلاک توسط یک عنصر گوک درجه دو نیز در شکل (۳) (ب) نمایش داده شده است. برای مشخص کردن یک مرکز درجه دو، علاوه بر اندیس نقطه شروع و اندیس نقطه پایان، از یک پارامتر دیگر نیز برای بیان میزان انحنا استفاده می‌شود. بنابراین، یک گوک درجه دو به‌صورت  $W(v_1, v_2, d, h_1, h_2)$  مشخص می‌شود که در آن عدد صحیح  $d$  پارامتری است که میزان انحراف منحنی را از یک پاره‌خط صاف مشخص می‌کند. منحنی مرکز درجه دو می‌تواند یک سهمی، هذلولی، یا بیضی را نمایش دهد؛ که استفاده از معادله سهمی رایج‌تر است.

تقریب گوک درجه دو برای یک بلاک از تصویر می‌تواند با رابطه زیر بیان شود:

(۴)

$$M_{W(v_1, v_2, d, h_1, h_2)}(x, y) = h_1 + (h_2 - h_1)W_{(v_1, v_2, d)}(x, y)$$

در لغت‌نامه گوک درجه دو تنها یک مجموعه محدود از مقادیر ممکن برای انحنا در نظر گرفته می‌شود. بر این اساس لغت‌نامه تبدیل گوک درجه‌دو، با توسعه لغت‌نامه گوک درجه‌یک، به شکل زیر تعریف می‌شود:

$$W = \{W_{(j,i,m,d)}: j = 0, \dots, \log_2 N; \quad (5)$$

$$i = 0, \dots, 4^i - 1; \\ m = 0, \dots, \left\lfloor W\left(\frac{N}{2^i}\right) \right\rfloor - 1; \\ d = -C, \dots, +C; \}$$

<sup>1</sup> second-order beamlet

که در آن سمبل  $d$  نشان‌دهنده میزان انحراف لبه درجه‌دو از مرکز درجه‌یک است، و مقادیر آن در این لغت‌نامه به مجموعه اعداد صحیح  $[-C, +C]$  محدود شده‌اند.

### ۴- تبدیل صفحک

دوربین‌های عمق مبتنی بر نور ساخت‌یافته بیشتر در محیط‌های داخلی مورد استفاده قرار می‌گیرند، که نور خورشید در فرایند سنجش عمق تداخلی ایجاد نمی‌کند. در این محیط‌ها اغلب اشیاء، ساخته دست بشر هستند و از سطوح صاف و صفحه‌گون تشکیل می‌شوند [19, 32]. توجه به این نکته در طراحی موجک هندسی مناسب برای بازنمایی تصاویر عمق صحنه‌های داخلی راه‌گشا خواهد بود. در این راستا، هر ناحیه از یک بلاک را می‌توان با یک صفحه توصیف کرد.

در تصاویر عمق مقدار هر پیکسل از تصویر فاصله یک نقطه از صحنه را تا صفحه دوربین نشان می‌دهد. اگر یک سطح صفحه‌گون در مقابل دوربین عمق قرار گیرد، مقادیر عمق پیکسل‌های متناظر با آن در تصویر عمق تغییرات غیرخطی خواهند داشت. برای درک ساده‌تر این موضوع، یک سطح صفحه‌گون را در مقابل دوربین در نظر بگیرید که بر محور  $Y$  در فضای سه‌بعدی  $(X, Y, Z)$  عمود، اما نسبت به محور  $X$  کمی چرخش یافته باشد. در این صورت، هنگامی که از نمای بالا به چنین سطحی بنگریم، نیم‌رخ آن مشابه شکل (۴) (الف) به‌صورت یک خط مشاهده خواهد شد. همان‌طور که در شکل (۴) (الف) نشان داده شده است، هر پیکسل از تصویر عمق فاصله یک نقطه از صحنه را تا دوربین سنجش می‌کند، و فاصله مرکز پیکسل‌ها بر روی صفحه حس‌گر دوربین یکسان است. در مقابل، در مورد صفحه نشان‌داده‌شده در صحنه سه بعدی، فاصله نقاط نمونه‌برداری شده توسط پیکسل‌های متوالی یکسان نبوده، و از سمت راست به سمت چپ به‌تدریج افزایش می‌یابد. به‌طور مشابه، مقادیر عمق نقاط نمونه‌برداری شده نیز مشابه شکل (۴) (ب) به‌تدریج از سمت راست به سمت چپ به‌صورت غیر خطی افزایش خواهد یافت. بنابراین در تصویر عمق، مقادیر عمق نقاط نمونه‌برداری شده از یک سطح صفحه‌گون، تغییرات غیرخطی و سهمی‌گون خواهند داشت. تمام نقاطی که در فضای سه‌بعدی صحنه بر روی یک صفحه قرار دارند، در معادله زیر صدق می‌کنند:

$$aX + bY + cZ + d = 0 \quad (6)$$

که در آن پارامترهای  $(a, b, c)$  جهت صفحه را تعیین کرده و پارامتر  $d$  متناظر با فاصله صفحه تا مبدأ مختصات است. از

به‌منظور تقریب تصاویر عمق، یک عنصر از لغت‌نامه مرکز می‌تواند با پارامترهای دو تابع گویا با رابطه (۱۰) تزیین شود تا هر یک از دو ناحیه حاصل را با یک صفحه توصیف کند. در این مقاله، چنین مرکز تزیین‌شده‌ای یک صفحک نامیده می‌شود. نمای سه‌بعدی سطح یک تابع گویا با رابطه (۱۰) در کنار یک عنصر صفحک درجه یک و یک عنصر صفحک درجه دو در شکل (۵) نمایش داده شده است.

تقریب صفحک درجه یک برای یک بلاک تصویر می‌تواند با رابطه زیر بیان شود:

$$\hat{f}(x, y) = \hat{p}_1(x, y) \cdot W_{(v_1, v_2)}(x, y) + \hat{p}_2(x, y) \cdot (1 - W_{(v_1, v_2)}(x, y)) \quad (11)$$

$$\hat{p}_i(x, y) = \frac{1}{d_i(x, y)} = \frac{1}{\alpha_i x + \beta_i y + \gamma_i} \quad (12)$$

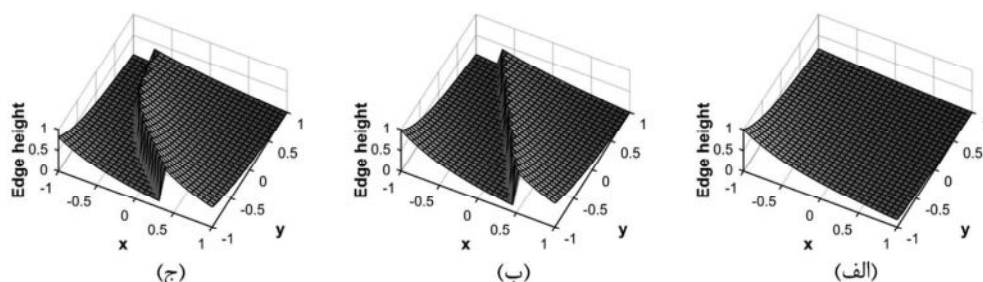
تقریب صفحک درجه دو یک بلاک نیز می‌تواند با رابطه‌ای مشابه و لحاظ کردن پارامتر انحنای لبه به شکل زیر مشخص گردد:

$$\hat{f}(x, y) = \hat{p}_1(x, y) \cdot W_{(v_1, v_2, d)}(x, y) + \hat{p}_2(x, y) \cdot (1 - W_{(v_1, v_2, d)}(x, y)) \quad (13)$$



(شکل-۴): (الف) طبیعت غیرخطی مقادیر عمق در تصویر عمق یک سطح صفحه‌گون: (الف) نحوه نمونه‌برداری از سطح صفحه‌گون توسط دوربین عمق، (ب) طبیعت غیرخطی مقادیر عمق بر روی صفحه تصویر

(Figure-4): Non-linearity of depth values in depth image of a planar surface (a) sampling mechanism of depth camera, (b) non-linearity of sampled depth values from palaner surface



(شکل-۵): نمای سه بعدی سطح مقادیر عمق بر روی شبکه پیکسل‌های تصویر حاصل از: (الف) یک تابع کسری خطی، (ب) صفحک درجه یک، (ج) صفحک درجه دو

(Figure-5): Surface plots of depth values on the image grid generated by (a) a linear-fractional function, (b) a first-order planelet, and (c) a second-order planelet

طرف دیگر، در یک دوربین عمق کالیبره‌شده، رابطه زیر بین مختصات  $(X, Y)$  یک نقطه و عمق سنجیده‌شده  $Z$  برای آن نقطه برقرار است [32]:

$$X = \frac{Z}{f}(x - c_x) \quad (7)$$

$$Y = \frac{Z}{f}(y - c_y) \quad (8)$$

که در آن سمبل  $f$  نشان‌دهنده فاصله کانونی دوربین عمق، سمبل‌های  $(c_x, c_y)$  نشان‌دهنده مکان نقطه اصلی در تصویر، و سمبل‌های  $(x, y)$  نشان‌دهنده مکان پیکسل مربوطه در تصویر است. با جایگذاری روابط (۷) و (۸) در رابطه (۶) کمی جابه‌جایی و به‌کارگرفتن سه نماد جدید  $\alpha$ ،  $\beta$ ، و  $\gamma$  به رابطه زیر برای نقاط یک ناحیه صفحه‌گون می‌رسیم:

$$\frac{1}{Z} = \alpha x + \beta y + \gamma \quad (9)$$

رابطه بالا می‌تواند به شکل یک تابع گویا<sup>۱</sup> بازنویسی شود، و نشان می‌دهد مقادیر عمق این نقاط بر روی تصویر عمق رابطه‌ای غیرخطی به شکل یک تابع گویا با مختصات مکانی خود دارند:

$$Z = \frac{1}{\alpha x + \beta y + \gamma} \quad (10)$$

<sup>۱</sup> یک تابع کسری که صورت و مخرج آن هر دو چند جمله‌ای هستند.

برای محاسبه پارامترهای صفحک هنگام تقریب یک بلاک، به عنوان یک رویکرد زیر بهینه اما سریع، پارامترهای لبه و پارامترهای تزئین می توانند مستقل از یکدیگر محاسبه شوند. از آنجایی که در تصاویر عمق لبه ها به طور عمومی تیز و قوی هستند، انتظار می رود که در بیش تر بلاک های تصویر تخمین تبدیل گوک از موقعیت لبه با تخمین تبدیل صفحک از موقعیت لبه یکسان باشد. بر این اساس، برای محاسبه سریع تر مقدار پارامترهای لبه می توان از محاسبه تبدیل گوک تصویر استفاده کرد [36]. در این صورت، برای محاسبه تبدیل صفحک تنها کافی است لبه پیدا شده توسط تبدیل گوک با دو تابع گویا مشابه رابطه (۱۰) تزئین شود.

محاسبه دقیق تبدیل گوک به کمک الگوریتم تبدیل گوک سنتی (WT) انجام می شود که شامل بررسی میزان خطای تمام گوک های ممکن در هر بلاک از تصویر است [14, 34]. این جستجوی کامل باید برای هر بلاک از درخت چهارگانه تصویر انجام شود. به همین دلیل، تبدیل گوک سنتی به زمان زیادی برای محاسبه تبدیل گوک تصویر نیاز دارد. به منظور افزایش سرعت محاسبه تبدیل گوک، الگوریتم های سریع اما زیر بهینه مختلفی ارائه شده اند؛ که همگی آن ها موازنه ای بین زمان و دقت محاسبه تبدیل گوک برقرار می سازند. از جمله مهم ترین الگوریتم های تبدیل گوک سریع می توان به شناسایی پیکسل های لبه و ردگیری لبه [24]، شناسایی پیکسل های لبه و برازش خط یا منحنی [25]، استنتاج گوک [16]، محاسبه ترتیبی گوک ها به کمک قضیه گرین<sup>۱</sup> [37]، و تبدیل گوک مبتنی بر گشتاور<sup>۲</sup> (MWT) [35] اشاره کرد. این الگوریتم ها، به طور عمومی به کاهش شدید دقت تبدیل گوک در بازنمایی تصویر منجر می شوند؛ تا جایی که نتایج حاصل از آن ها در بسیاری از کاربردها از نتایج روش های غیر هندسی ضعیف تر خواهد بود. بر همین اساس، به منظور ارائه موازنه بهتری میان زمان و دقت محاسبه تبدیل گوک، الگوریتم تبدیل گوک تکراری درجه یک<sup>۳</sup> (IWT) [38] و الگوریتم تبدیل گوک تکراری درجه دو<sup>۴</sup> (ISWT) [39] معرفی شده اند که برای تخمین پارامترهای گوک از روش کمترین مربعات خطای غیرخطی استفاده می کنند. در این مقاله، برای محاسبه پارامترهای لبه در هر بلاک از الگوریتم تبدیل گوک تکراری درجه دو (ISWT) استفاده می شود؛ سپس به منظور محاسبه تبدیل صفحک تصویر، برای تقریب هر یک از دو

ناحیه در هر بلاک از برازش یک تابع گویا با رابطه (۱۰) استفاده می شود.

### ۱-۵- تبدیل گوک تکراری درجه دو

یک روش رایج برای محاسبه پارامترهای یک مدل بر اساس داده های موجود، استفاده از روش کمترین مربعات خطا است که برای توابع خطای مشتق پذیر می تواند مورد استفاده قرار گیرد؛ اما، از آنجایی که تابع گوک  $W_{(v_1, v_2)}$  در رابطه (۱) در محل لبه دارای ناپیوستگی بوده و یک تابع مشتق پذیر نیست، روش بهینه سازی کمترین مربعات خطا نمی تواند به طور مستقیم برای محاسبه پارامترهای آن مورد استفاده قرار گیرد. با این وجود، لبه تیز در تعریف گوک می تواند با یک لبه هموار جایگزین شود تا امکان استفاده از کمترین مربعات خطا برای محاسبه پارامترهای آن فراهم آید [39].

برای سادگی بیان، دستگاه مختصات چرخش یافته و انتقال یافته  $(\tau, s)$  را که با چرخش دستگاه مختصات  $(x, y)$  به اندازه زاویه  $\theta$  در جهت خلاف عقربه های ساعت (CCW) و انتقال دستگاه حاصل به اندازه  $t$  در طول محور  $\tau$  تولید شده است، در نظر می گیریم. در این صورت، رابطه دستگاه مختصات چرخش و انتقال یافته  $(\tau, s)$  با دستگاه مختصات اصلی  $(x, y)$  به شکل زیر خواهد بود:

$$\begin{bmatrix} \tau \\ s \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} t \\ 0 \end{bmatrix} \quad (14)$$

در تعریف گوک تسهیل شده برای مدل سازی نیم رخ<sup>۵</sup> لبه به جای یک تابع دو ضابطه ای از تابع هموار و مشتق پذیر سیگموئید به شکل زیر استفاده می شود:

$$\mathcal{RW}_{(\theta, t, \sigma)}(x, y) = S\left(\frac{\tau}{\sigma}\right) \quad (15)$$

که در آن برای بازنمایی مرکز درجه یک از دو پارامتر  $(\theta, t)$  استفاده شده است، و تابع  $S(x) = 1/(1 + e^{-x})$  همان تابع سیگموئید است. پارامتر  $\theta$  زاویه و پارامتر  $t$  اندازه خطی را نشان می دهد که بر لبه عمود است و از مبدأ مختصات بلاک در مرکز آن می گذرد. پارامتر  $\sigma$  میزان تیز بودن لبه را کنترل می کند که در آزمایش های ما با مقدار ثابت  $\sigma = 0.5$  پیکسل مقداره ای شده است. از این رویکرد می توان برای تعریف مدل مشتق پذیر گوک تسهیل شده درجه دو نیز استفاده کرد:

$$\mathcal{RW}_{(\theta, t, \eta, \sigma)}(x, y) = S\left(\frac{\tau - \eta s^2}{\sigma}\right) \quad (16)$$

که در آن پارامتر جدید  $\eta$  میزان انحنای لبه را تعیین می کند.

<sup>۵</sup> profile

<sup>۱</sup> Green's theorem

<sup>۲</sup> moments-based wedgelet transform (MWT)

<sup>۳</sup> iterative wedgelet transform (IWT)

<sup>۴</sup> iterative second-order wedgelet transform (ISWT)

تخمین تکراری پارامترهای گوک درجه‌دو از یک مرکز اولیه آغاز می‌شود. هر چه این مرکز اولیه به مرکز واقعی نزدیک‌تر باشد، الگوریتم در تعداد تکرارهای کمتری هم‌گرا خواهد شد؛ به همین دلیل، در این مقاله، به‌منظور افزایش سرعت هم‌گرایی در تخمین تکراری پارامترهای گوک درجه‌دو از نتیجه تبدیل گوک درجه‌یک مبتنی بر گشتاور (MWT) [35] به‌عنوان مقادیر آغازین پارامترها در فرآیند بهینه‌سازی استفاده شد.

به‌منظور تولید بازنمایی گوک درجه‌دو برای یک تصویر ورودی، پس از محاسبه پارامترهای گوک درجه‌دو در دامنه پیوسته، باید این پارامترها به مقادیر گسسته تبدیل شده و در درخت چهارگانه ذخیره شوند. به این منظور، برای مرکز درجه‌یک متناظر با خط مشخص شده توسط مقادیر  $(\theta, t)$ ، اندیس نقاط ابتدا و انتهای مرکز  $b(v_1, v_2)$  با بررسی محل تقاطع خط یادشده با چهار ضلع بلاک تعیین شد؛ سپس، برای بازنمایی میزان انحراف مرکز درجه دو از مرکز درجه‌یک به کمک مقادیر گسسته، مقدار پارامتر  $d$  برای مرکز گسسته درجه دو  $b(v_1, v_2, d)$  به‌صورت  $d = \eta \frac{n^2}{4}$  محاسبه شد (که  $n$  در آن نشان‌دهنده عرض بلاک است).

از جنبه نظری، پیچیدگی محاسباتی تبدیل گوک تکراری درجه دو (ISWT) برای محاسبه تبدیل گوک یک تصویر  $N \times N$  از مرتبه  $O(N^2 \log_2 N)$  است. بر این اساس، تبدیل گوک تکراری درجه‌دو (ISWT) به کمینه پیچیدگی زمانی ممکن برای محاسبه تبدیل گوک درجه‌دو دست یافته است.

## ۲-۵- تزئین لبه با توابع گویا

پس از محاسبه پارامترهای لبه به کمک تبدیل گوک تکراری درجه دو (ISWT)، به‌منظور محاسبه تبدیل صفحک، هر یک از دو ناحیه حاصل در هر بلاک تصویر باید با یک تابع گویا به شکل رابطه (۱۰) تقریب زده شود. بدین منظور، پارامترهای دو تابع یادشده می‌توانند با کمینه‌سازی تابع هزینه زیر محاسبه شوند:

$$D = \sum_{(x_i, y_i) \in R} \left( \frac{1}{\alpha_1 x + \beta_1 y + \gamma_1} - I(x_i, y_i) \right)^2 + \sum_{(x_i, y_i) \in \bar{R}} \left( \frac{1}{\alpha_2 x + \beta_2 y + \gamma_2} - I(x_i, y_i) \right)^2 \quad (17)$$

که در آن سمبل‌های  $R$  و  $\bar{R}$  دو ناحیه حاصل از لبه محاسبه شده  $b(v_1, v_2, d)$  را نشان می‌دهند. نحوه کمینه‌سازی این رابطه در پیوست ۱ شرح داده شده است.

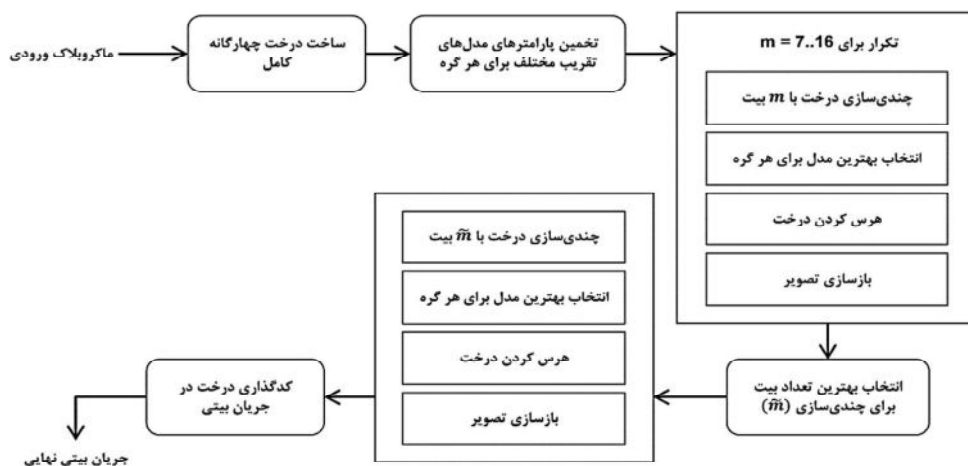
با جایگزین کردن گوک تسهیل‌شده به‌جای گوک سنتی در تابع تقریب گوک، تابع خطای تقریب که باید برای هر بلاک کمینه‌سازی شود به یک تابع مشتق‌پذیر تبدیل می‌شود؛ اما این تابع نسبت به پارامترهای  $(\theta, t, \eta)$  غیرخطی و نسبت به پارامترهای  $(h_1, h_2)$  خطی است. به همین دلیل، تخمین پارامترهای موقعیت لبه یعنی  $(\theta, t, \eta)$  باید با روش کمترین مربعات خطای غیرخطی و به‌صورت تکراری انجام شود. به ازای هر مقدار مشخص برای  $(\theta, t, \eta)$ ، چنانچه مقادیر  $\theta$  و  $t$  و  $\eta$  به‌عنوان ثابت در نظر گرفته شوند، مقدار بهینه پارامترهای تزئین  $(h_1, h_2)$  می‌تواند به شکل بسته محاسبه شود. الگوریتم نمایش داده شده در شکل (۶) نحوه محاسبه تکراری پارامترهای گوک درجه‌دو را تشریح می‌کند.

ورودی: I بلاک تصویر، $\theta_0$ چرخش اولیه، $t_0$ انتقال اولیه خروجی: $(\theta^*, t^*, \eta^*, h_1^*, h_2^*)$ پارامترهای گوک تسهیل شده، $D^*$ خطای گوک تسهیل شده	
۱.	$D^* \leftarrow \infty$
۲.	$\sigma \leftarrow 0.5, \eta \leftarrow 0, t \leftarrow t_0, \theta \leftarrow \theta_0$
۳.	تجزیه بلاک به دو ناحیه به کمک گوک $RW_{(\theta, t, \eta)}$
۴.	محاسبه ضرایب تزئین $h_1$ و $h_2$ به کمک محاسبه میانگین هر ناحیه حاصل
۵.	بازسازی تقریب بلاک به کمک گوک تزئین شده
۶.	محاسبه تفاضل بلاک بازسازی شده با بلاک ورودی
۷.	محاسبه خطای $D$ میان بلاک بازسازی شده و بلاک ورودی با معیار SSE
۸.	$n \leftarrow 0$
۹.	تا زمانی که $(D < D^*)$ و $(n < MaxIter)$ مراحل زیر را تکرار کن
۹.۱.	$n \leftarrow n + 1$
۹.۲.	$D^* \leftarrow D, \eta^* \leftarrow \eta, t^* \leftarrow t, \theta^* \leftarrow \theta$
۹.۳.	محاسبه ژاکوبین تابع خطا
۹.۴.	محاسبه اندازه گام $\Delta \xi$ بر اساس ژاکوبین و خطا
۹.۵.	به‌روزرسانی پارامترهای لبه $(\theta, t, \eta)$ به‌صورت $\xi \leftarrow \xi + \Delta \xi$
۹.۶.	تجزیه بلاک به دو ناحیه به کمک گوک $RW_{(\theta, t, \eta)}$
۹.۷.	محاسبه ضرایب تزئین $h_1$ و $h_2$ به کمک محاسبه میانگین هر ناحیه حاصل
۹.۸.	بازسازی تقریب بلاک به کمک گوک تزئین شده
۹.۹.	محاسبه تفاضل بلاک بازسازی شده با بلاک ورودی
۹.۱۰.	محاسبه خطای $D$ میان بلاک بازسازی شده و بلاک ورودی با معیار SSE

(شکل-۶): الگوریتم تخمین تکراری پارامترهای گوک درجه دو

برای یک بلاک از تصویر در تبدیل گوک تکراری درجه دو

(Figure-6): Iterative parameters estimation algorithm for second-order wedgelet approximation of an image block in iterative second-order wedgelet transform



(شکل-۷): ساختار روش پیشنهادی برای فشرده‌سازی یک ماکرو بلاک از تصویر عمق  
(Figure-7): Structure of the proposed depth-image compression scheme

نابه‌جایی استفاده شده است [9, 32, 34, 39, 40]. در این مقاله، برای فشرده‌سازی هر تصویر ابتدا تصویر به ماکرو بلاک‌هایی با اندازه ثابت  $64 \times 64$  تقسیم و سپس مراحل نشان داده‌شده در شکل (۷) برای هر ماکرو بلاک انجام شد.

### ۶-۱- تقریب و کدگذاری بلاک‌های تصویر

در روش پیشنهادی این مقاله برای فشرده‌سازی تصاویر عمق، هر بلاک توسط یکی از چهار تابع زیر بازنمایی می‌شود:

- تابع ثابت: بلاک را به‌صورت یک ناحیه با عمق یکنواخت تقریب می‌زند.
- تابع گویا: مقادیر عمق پیکسل‌های بلاک را به‌کمک یک تابع گویا تقریب می‌زند.
- تابع گوک درجه‌دو: از یک مرکز درجه‌دو برای ناحیه‌بندی بلاک به دو ناحیه استفاده کرده و سپس هر ناحیه را با عمق یکنواخت تقریب می‌زند.
- تابع صفحک درجه‌دو: از یک مرکز درجه‌دو برای ناحیه‌بندی بلاک به دو ناحیه استفاده کرده و سپس هر ناحیه را با یک تابع گویا تقریب می‌زند.

هر یک از چهار عنصر بالا به‌منظور کدگذاری در جریان بیتی به بودجه بیتی متفاوتی نیاز دارند. در یک بلاک  $M \times M$  از تصویر، برای هر گوک غیر تباهیده، اندیس هر یک از نقاط پایانی بین 0 تا  $4M - 1$  است. بنابراین، برای کدگذاری اندیس هر نقطه پایانی به  $r = \log_2(4M)$  بیت نیاز خواهیم داشت. به‌طور معمول در بازنمایی یک تصویر، برای تعداد قابل توجهی از بلاک‌ها مقدار انحنای لبه برابر با صفر است. بنابراین، برای صرفه‌جویی در کدگذاری بلاک‌ها به کمک یک گوک یا صفحک، صفر بودن یا غیر صفر بودن مقدار انحنای لبه می‌تواند با

## ۶- کدگذاری تصاویر عمق به کمک موجک‌های هندسی

در بخش قبل نحوه محاسبه سریع تبدیل صفحک تصویر تشریح شد. در این بخش، تبدیل صفحک محاسبه‌شده برای تصویر به‌منظور تولید یک بازنمایی تنک از تصویر با بالاترین کیفیت در یک نرخ بیت محدود مورد استفاده قرار خواهد گرفت. یک رویکرد رایج برای دست‌یابی به بیشترین کیفیت و کمترین نرخ بیت در فشرده‌سازی تصاویر تقسیم تصویر به بلاک‌هایی با اندازه‌های مختلف و استفاده از یک چارچوب مبتنی بر بهینه‌سازی نرخ خرابی است. در این رویکرد، تصویر به کمک درخت چهارگانه به بلاک‌هایی با اندازه‌های مختلف تقسیم می‌شود. گره ریشه درخت چهارگانه کل تصویر را پوشش می‌دهد؛ سپس در هر سطح، بلاک قبلی به چهار بلاک هم‌اندازه تقسیم می‌شود. این تقسیمات متوالی تا رسیدن به بلاک‌های  $1 \times 1$  ادامه پیدا می‌کند. گره‌های هر سطح از درخت چهارگانه کل تصویر را پوشش می‌دهند. به این ترتیب، درخت چهارگانه یک بازنمایی افزونه از تصویر را تولید می‌کند.

برای بازسازی تصویر به تمام گره‌های درخت چهارگانه نیازی نیست، بلکه زیرمجموعه‌ای از گره‌ها که کل تصویر را پوشش دهند، کافی خواهد بود. برای انتخاب بهترین گره‌ها جهت تقریب تصویر، هر بلاک از درخت چهارگانه با مدل‌های مختلف تقریب‌زده‌شده، پارامترها چندبندی‌سازی می‌شوند، و سپس درخت بر اساس تابع هزینه نرخ-خرابی هرس می‌شود. در نهایت، بهترین درخت هرس‌شده در فایل کدگذاری می‌شود. از این رویکرد در قبل برای دست‌یابی به کمترین مقدار خرابی در هر نرخ بیت حین فشرده‌سازی تصاویر خاکستری و تصاویر

از آنجایی که پیچیدگی ساختاری بخش‌های مختلف تصویر با یکدیگر یکسان نیست، بهتر است الگوریتم فشرده‌سازی تصویر برای دست‌یابی به دقت بالاتر در بازنمایی بخش‌های پیچیده‌تر از بلاک‌های کوچک‌تر و بیشتری استفاده کند. در مقابل، به‌منظور کاهش نرخ بیت مصرفی برای بخش‌های ساده‌تر تصویر باید از بلاک‌های بزرگ‌تر و کم‌تری استفاده شود. بدین منظور، درخت چهارگانه می‌تواند بر اساس تابع هزینه نرخ-خرابی معادله (۱۸) از پایین به بالا هرس شود، تا در نهایت گره‌های برگ درخت هرس‌شده برای بازنمایی تصویر مورد استفاده قرار گیرند [9, 40].

هرس درخت چهارگانه با تکیه بر این موضوع انجام می‌شود که در هر زیر درخت از درخت چهارگانه تنها گره‌های برگ در تقریب تصویر شرکت می‌کنند. طی فرآیند هرس، در مورد هر گره میانی، چنانچه هزینه بازنمایی تصویر به کمک برگ‌های زیر درخت آن گره از هزینه بازنمایی تصویر به کمک خود آن گره بیشتر باشد، فرزندان آن گره هرس می‌شوند. بنابراین، شرط هرس برای هر گره میانی  $N$  عبارت است از

$$J(N) < \sum_{L \in \text{Leaves}(N)} J(L) + \lambda \sum_{M \in \text{Internals}(N)} R(M) \quad (19)$$

که در آن  $\text{Leaves}(N)$  نشان‌دهنده مجموعه گره‌های برگ زیردرختی با ریشه  $N$ ، و  $\text{Internals}(N)$  نماینده مجموعه گره‌های میانی این زیر درخت است.

تعداد بیت مصرفی برای کدگذاری هر مقدار عمق که با سمبل  $m$  نشان داده شد، می‌تواند به‌عنوان یک مقدار ثابت قبل از فرآیند فشرده‌سازی تعیین شود؛ اما، به‌طور معمول برای دستیابی به کارایی بالاتر در فشرده‌سازی، مقدار بهینه این پارامتر با انجام یک جستجوی خطی بر روی مجموعه مقادیر محتمل تعیین می‌شود. مقدار بهینه پارامتر  $m$  برای تصاویر عمق، می‌تواند به شکل زیر تعیین شود [9]:

$$m^* = \min_m D_m + \lambda R_m \quad (20)$$

الگوریتم نهایی برای فشرده‌سازی تصویر به کمک موجک‌های هندسی در شکل (۸) خلاصه شده است.

## ۷- افزایش کارایی کدگذار با پیش‌بینی مکانی

در افراز نهایی تولیدشده برای هر تصویر، به‌طور معمول بلاک‌های هم‌جوار از شباهت و همبستگی زیادی برخوردار بوده، و دست‌یابی به نرخ فشرده‌سازی بهتر با پیش‌بینی بعضی بلاک‌ها از روی بلاک‌های همسایه نیز امکان‌پذیر است. بر

ارسال یک بیت با مقدار 0 یا 1 مشخص شود. در نهایت از آنجایی که مقادیر انحنا در دامنه گسسته  $[-C..+C]$  قرار دارند، هر مقدار انحنا غیر صفر می‌تواند با  $p = 1 + \log_2 C$  بیت کدگذاری شود. برای کدگذاری هر تابع گویا نیز از مقدار عمق آن برای سه گوشه مشخص از بلاک استفاده می‌شود، که به کدگذاری سه مقدار عمق نیاز دارد. بر این اساس، چنانچه از  $m$  بیت برای کدگذاری هر مقدار عمق استفاده شود، هزینه بیتی هر تابع تقریب استفاده شده به شکل زیر خواهد بود:

• تابع ثابت:  $2 + m$  بیت

• تابع گویا:  $2 + 3m$  بیت

• تابع گوک:  $2 + 2r + 2m + 1 + p$  بیت برای انحنا غیر

صفر، و  $2 + 2r + 2m + 1$  بیت برای انحنا صفر

• تابع صفحک:  $2 + 2r + 6m + 1 + p$  بیت برای انحنا غیر صفر، و  $2 + 2r + 6m + 1$  بیت برای انحنا صفر

در مورد بلاک‌های  $2 \times 2$ ، تعداد گوک‌های ممکن برابر با شش خواهد بود. به همین دلیل، به‌جای ارسال اندیس نقاط پایانی گوک، بهتر است، اندیس گوک در لغت‌نامه بلاک‌های  $2 \times 2$  ارسال شود؛ که تنها به سه بیت نیاز خواهد داشت.

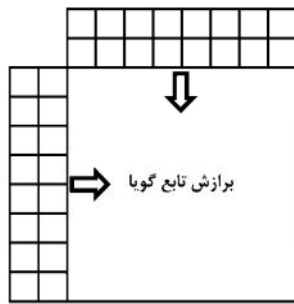
برای تصمیم‌گیری در مورد اینکه هر بلاک با کدام تابع تقریب زده شود، تابع هزینه نرخ-خرابی زیر برای هر چهار تابع تقریب بالا ارزیابی شده و تقریبی مورد استفاده قرار می‌گیرد که هزینه کمتری داشته باشد [9]:

$$J = D + \lambda R \quad (18)$$

که در آن سمبل خرابی  $D$  میزان خطای تقریب تصویر را نشان می‌دهد (که با معیار خطای SSE محاسبه می‌شود)، سمبل نرخ  $R$  تعداد بیت مصرفی و ضریب  $\lambda$  میزان اهمیت استفاده از نرخ پایین‌تر را نشان می‌دهد. هر چه مقدار  $\lambda$  بزرگ‌تر باشد، در فشرده‌سازی تصویر از تعداد بیت کمتری استفاده می‌شود.

## ۲-۶- تولید بازنمایی بهینه با هرس درخت

در تبدیلات مختلف خانواده موجک‌های هندسی و فقی، درخت چهارگانه یک بازنمایی افزونه از تصویر را تولید می‌کند که هر سطح آن تقریبی از تصویر را با سطحی از دقت فراهم می‌آورد. تقریب تولیدشده توسط سطوح پایین‌تر درخت، نسبت به تقریب تولیدشده توسط سطوح بالاتر درخت از دقت بیشتری برخوردار است. از طرف دیگر، هر چه درخت چهارگانه پایین‌تر می‌رویم، تعداد بلاک‌ها نیز افزایش می‌یابد. بنابراین، برای تولید یک تقریب دلخواه از تصویر، می‌توان تنها زیرمجموعه‌ای از گره‌های درخت چهارگانه را برای بازنمایی تصویر به کار برد؛ طوری که تمام تصویر پوشش داده شده و بلاک‌های انتخاب‌شده فاقد هم‌پوشانی باشند.



(شکل-۹): پیش‌بینی بین بلاکی افقی و عمودی در کدگذار مبتنی بر صفحک

(Figure-9): Horizontal and vertical inter-block prediction modes of planelets based depth image coder

## ۸- نتایج

در این بخش عملکرد کدگذار مبتنی بر صفحک در فشرده‌سازی تصاویر عمق به کمک آزمایش‌های مختلف ارزیابی شده و با موجک‌های هندسی و فقی پیشین، و برخی کدگذارهای مدرن تصاویر عمومی مقایسه می‌شود. در گروه موجک‌های هندسی، عملکرد روش پیشنهادی با کدگذار مبتنی بر تبدیل گوک و کدگذار مبتنی بر تبدیل شیبک مقایسه خواهد شد. در میان کدگذارهای مدرن نیز عملکرد الگوریتم پیشنهادی در مقابل عملکرد کدگذارهای JMK1a و H.264/AVC 14-bit و JPEG2000 16-bit ارزیابی خواهد شد. در آزمایش‌های این بخش تمام روش‌های پیش‌بینی بین بلاکی برای کدگذار H.264 در حالت فعال قرار داده شدند.

به‌منظور ارزیابی عملکرد کدگذارهای مختلف در فشرده‌سازی تصاویر عمق، دو مجموعه از تصاویر عمق مصنوعی و واقعی در آزمایش‌ها مورد استفاده قرار گرفتند. مجموعه تصاویر عمق مصنوعی شامل سی تصویر فاقد نوفه است که از یک آپارتمان و یک دفتر کار به کمک شبیه‌سازی رایانه‌ای تهیه شده‌اند. این تصاویر طوری از هشت جریان پایگاه تصاویر ICL-NUIM RGBD Benchmark انتخاب شده‌اند که محتوای آن‌ها تا حد ممکن متفاوت باشد [41]. مجموعه تصاویر عمق واقعی نیز شامل سی تصویر عمق است که به کمک دوربین کینکت تهیه شده و طوری از پایگاه تصاویر Washington RGBD Scenes V2 انتخاب شدند که صحنه‌های متفاوتی از محیط‌های داخلی در دنیای واقعی را پوشش دهند [42].

به‌منظور ارزیابی عملکرد روش‌های مختلف از نظر کیفیت تصاویر خروجی از ارزیابی عددی و بصری کیفیت تصاویر استفاده شد. به‌علاوه، زمان مصرفی الگوریتم‌های مختلف نیز بر روی یک دستگاه رایانه‌ای مشترک ارزیابی شد.

همین اساس، در روش پیشنهادی این مقاله، از دو حالت پیش‌بینی افقی و پیش‌بینی عمودی نیز برای کدگذاری هر بلاک استفاده شد. در این دو حالت، مشابه آنچه در شکل (۹) نشان داده شده است، یک تابع گویا با رابطه (۱۰) بر روی پیکسل‌های دو ستون هم‌جوار یا دو سطر هم‌جوار بلاک جاری برازش و سپس از تابع حاصل برای پیش‌بینی مقادیر عمق پیکسل‌های بلاک جاری استفاده می‌شود.

برای استفاده از این روش‌های پیش‌بینی مکانی در الگوریتم پیشنهادی این مقاله، پس از هر بار هرس درخت، حالت‌های پیش‌بینی نیز بررسی شده و تصمیم بهینه بین پیش‌بینی افقی، پیش‌بینی عمودی، و بازنمایی به کمک توابع تقریب به کمک تابع هزینه نرخ خرابی تعیین شد. برای اطمینان از این‌که قبل از هر بلاک، تمامی بلاک‌های بالا و سمت چپ آن بازسازی شده‌اند، بلاک‌های درخت چهارگانه با کمک یک پیمایش نخست عمق با ترتیب زیر درخت بالا-چپ، زیر درخت بالا-راست، زیر درخت پایین-چپ، و زیر درخت پایین-راست بازسازی شدند. در فرآیند چندی‌سازی ضرایب و تعیین بهترین مدل تقریب برای هر بلاک نیز از همین ترتیب پیمایش استفاده شد.

ورودی: I تصویر ورودی، λ ارزش نرخ بیت	
خروجی: فایل جریان بیتی	
۱.	ساخت درخت چهارگانه کامل
۲.	محاسبه تبدیل گوک تصویر به کمک درخت چهارگانه
۳.	محاسبه تبدیل صفحک تصویر به کمک درخت چهارگانه
۴.	$m^* \leftarrow J, J^* \leftarrow \infty$
۵.	تکرار مراحل زیر برای هر مقدار m از 7 تا 16
۵.۱.	چندی‌سازی کلیه گره‌های درخت با m بیت
۵.۲.	تعیین بهترین تابع تقریب در هر گره
۵.۳.	هرس درخت حاصل
۵.۴.	محاسبه هزینه کل درخت $J \leftarrow D + \lambda R$
۵.۵.	اگر $(J < J^*)$ آنگاه
۵.۵.۱.	$J^* \leftarrow J$
۵.۵.۲.	$m^* \leftarrow m$
۶.	چندی‌سازی کلیه گره‌های درخت با $m^*$ بیت
۷.	تعیین بهترین تابع تقریب در هر گره
۸.	هرس درخت حاصل
کدگذاری ساختار درخت و ضرایب گره‌های برگ در جریان بیتی	

(شکل-۸): الگوریتم فشرده‌سازی تصویر به کمک

موجک‌های هندسی

(Figure-8): Algorithm of image compression using geometrical wavelets

## ۸-۱- ارزیابی بصری کیفیت تصویر

ارزیابی بصری کیفیت تصاویر بازسازی‌شده امکان شناسایی انواع تخریب‌های تولیدشده توسط الگوریتم‌های فشرده‌سازی مختلف را فراهم می‌سازد. به‌علاوه، زمانی که با ویدئوهای نوفه‌ای سر و کار داریم، ارزیابی بصری امکان بررسی میزان مقاومت الگوریتم‌ها در برابر نوفهٔ دنیای واقعی و توانایی آن‌ها در بازیابی ساختارهای هندسی اصلی صحنه را فراهم می‌کند. به‌منظور ارزیابی تأثیر استفاده از مدل تقریب گویا به‌جای مدل‌های ثابت و خطی، در نخستین آزمایش دو تصویر عمق نشان داده شده در شکل (۱۰) که متشکل از چند صفحه اصلی هستند در نرخ بیت بسیار پائین  $0.2 \text{ bpp}$  توسط کدگذارهای مبتنی بر گوک، شیبک، صفحک، و کدگذارهای مدرن عمومی فشرده‌سازی شدند. تصویر عمق بازسازی‌شده به‌کمک هر یک از کدگذارهای یادشده با تصویر رنگ متناظرش ترکیب شد و پس از نگاشت نقاط به فضای سه‌بعدی، صحنه سه‌بعدی حاصل برای تصویر عمق مصنوعی سطر نخست شکل (۱۰) در شکل (۱۱) و برای تصویر عمق واقعی سطر دوم شکل (۱۰) در شکل (۱۲) نمایش داده شد. همان‌طور که در این دو شکل مشاهده می‌شود، در نرخ بیت بسیار پایین، کدگذار مبتنی بر گوک که نتیجه آن با برجسب Wedgelets مشخص شده، هر صفحه را به‌صورت پله‌پله و با کمک قطعاتی با عمق ثابت بازسازی کرده است. این در حالی است که کدگذار مبتنی بر شیبک که نتیجه آن با برجسب Platelets مشخص شده است، با بهره‌گیری از مدل تقریب خطی به بازسازی بسیار بهتری از هر صفحه منجر شده است؛ اما به‌دلیل عدم سازگاری مدل استفاده شده در تبدیل شیبک با تغییرات واقعی عمق در نواحی صفحه‌گون، هر صفحه در مرز بلاک‌ها دچار شکستگی شده است. کدگذار مبتنی بر صفحک که نتیجه آن با برجسب Planelets مشخص شده، به‌دلیل استفاده از مدل غیرخطی منطبق با طبیعت تغییرات عمق در نواحی صفحه‌گون، به بازسازی بهتری برای هر صفحه منجر شده است؛ تا جایی که در مرز بلاک‌ها نیز پیوستگی صفحه حفظ شده و از شکستگی اجتناب شده است. به‌علاوه، استفاده از روش‌های پیش‌بینی بین بلاکی در کدگذار مبتنی بر صفحک که نتیجه آن با برجسب  $\text{Planelets} + \text{SP}$  مشخص شده است، باعث تقریب بهتر کنج‌ها می‌شود.

در گروه کدگذارهای مدرن تصاویر عمومی، کدگذار JPEG2000 به‌شدت صفحات را تخریب کرده و باعث موج‌دارشدن سطوح و هموارشدن کنج‌ها شده است. درنهایت، کدگذار H.264 با بهره‌گیری از روش‌ها کدگذاری کارآمد از بلاک‌های کوچک زیادی برای بازسازی تصویر بهره گرفته، اما

تولید نقاط پرت در اطراف کنج‌ها و ایجاد ناپیوستگی در مرز بلاک‌ها در نتایج آن مشهود است. این در حالی است که کدگذار مبتنی بر صفحک مجهز به پیش‌بینی بین بلاکی، عملکرد بسیار خوبی را در حفظ صفحات و کنج‌ها داشته است. به‌منظور مقایسه کدگذارهای هندسی و مدرن در بازسازی لبه‌های تیز در نرخ بیت مشابه، در دومین آزمایش، یک تصویر عمق مصنوعی و یک تصویر عمق واقعی توسط کدگذارهای مختلف در نرخ بیت  $0.3 \text{ bpp}$  فشرده شدند. نسخهٔ بازسازی‌شده یک لبه تیز در هر یک از این تصاویر برای کدگذارهای مختلف در شکل‌های (۱۳) و (۱۴) نمایش داده شده است. در نرخ بیت بسیار پایین، کدگذار عمق مبتنی بر تبدیل گوک، به‌دلیل مدل‌سازی کلیه پیکسل‌های هر ناحیه هموار عمق با یک مقدار عمق ثابت، لبه‌های تیز را به‌صورت دندانه‌دار و نواحی هموار را به‌صورت پله‌پله بازسازی می‌کند. در مقابل، کدگذارهای عمق مبتنی بر شیبک و صفحک، با بهره‌گیری از مدل‌های تقریب پیچیده‌تر، به حفظ بهتر لبه‌ها و سطوح منجر می‌شوند؛ درنهایت، استفاده از روش‌های پیش‌بینی بین بلاکی در کدگذار مبتنی بر صفحک نیز باعث تقریب بهتر لبه‌های تیز تصویر با بهره‌برداری بهتر از بودجه بیتی می‌شود.

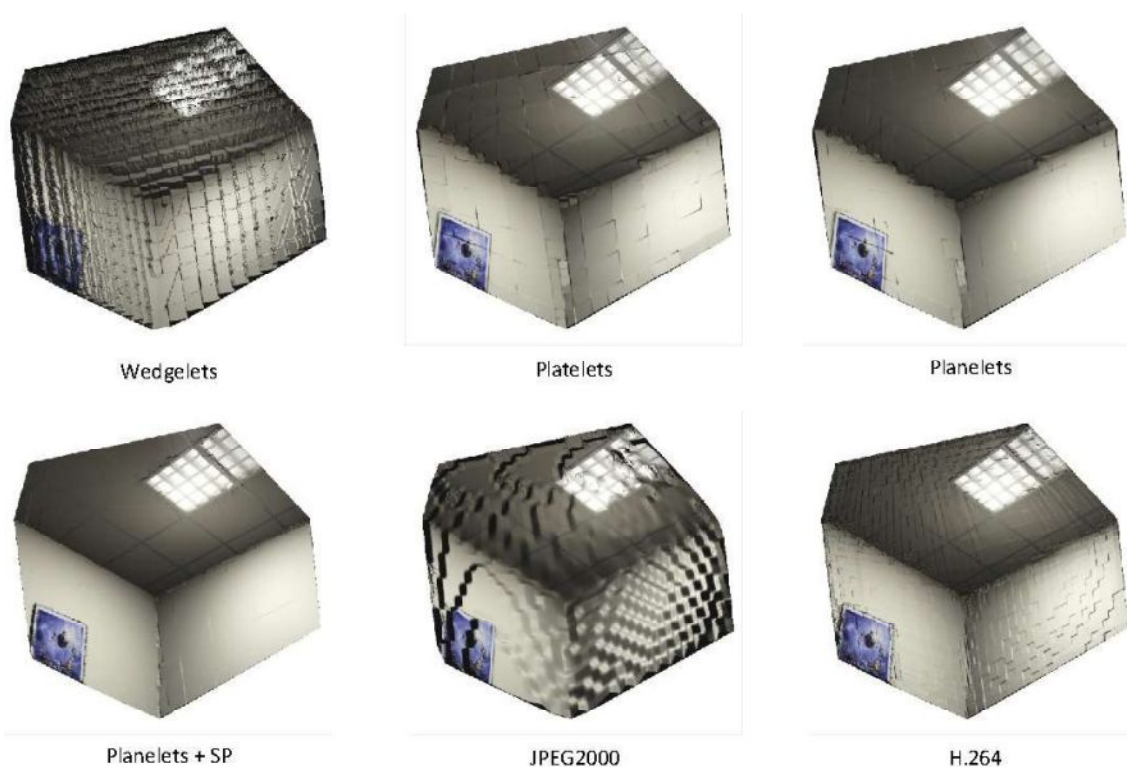
در میان کدگذارهای مدرن تصاویر عمومی، عملکرد کدگذار JPEG2000 در بازنمایی لبه‌ها بسیار ضعیف بوده و باعث هموارشدن لبه<sup>۱</sup> و موج‌دارشدن<sup>۲</sup> نواحی اطراف آن در شکل‌های (۱۳) و (۱۴) شده است. در مقابل، کدگذار H.264 عملکرد قابل قبولی در حفظ نواحی هموار داشته است؛ اما باعث موج‌دارشدن نواحی اطراف لبه نیز می‌شود. هر دو کدگذار یادشده در کنار لبه‌های تیز نقاطی غیرواقعی تولید کرده، و موجب تخریب لبه و نواحی اطراف آن می‌شوند. در مقابل، کدگذار مبتنی بر صفحک تا حد ممکن از هموارسازی لبه‌های تیز خودداری می‌کند.

هنگامی که با تصاویر عمق نوفه‌ای مواجه هستیم، همان‌طور که در شکل (۱۴) نشان داده شده، برتری موجک‌های هندسی در فیلترکردن نوفه و بازیابی لبه‌های تیز واقعی حین فشرده‌سازی داده‌های عمق نوفه‌ای قابل مشاهده است. با وجود نویز در تصاویر ورودی از دوربین کینکت، کدگذار مبتنی بر صفحک نه‌تنها به‌خوبی لبه‌ها و نواحی هموار را بازسازی کرده، بلکه بخشی از نویز در تصویر ورودی را نیز فیلتر کرده است. این در حالی است که کدگذار JPEG2000 نوفه را تشدید کرده و هم لبه‌ها و هم سطوح را به‌شدت تخریب کرده است. کدگذار H.264 نیز در اطراف لبه‌ها باعث تولید نقاط غیرواقعی شده و با هموارکردن لبه‌ها موجب محوشدن مرز برخی اشیاء شده است.

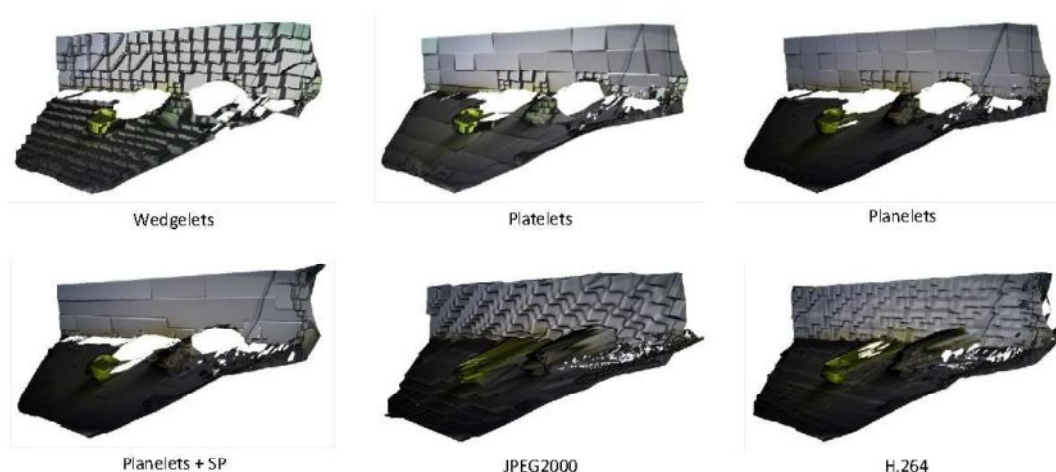
<sup>1</sup>blurring artifact<sup>2</sup>ringing artifact



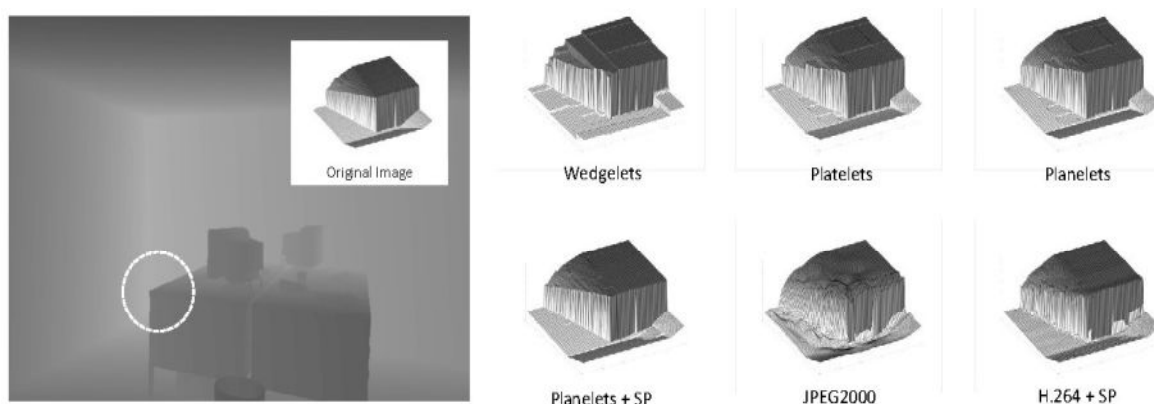
(شکل-۱۰): دو تصویر عمق صفحه‌گون مصنوعی و واقعی که به صورت سه‌بعدی نمایش داده شده‌اند  
(Figure-10): Sample synthetic and real-world depth images and their corresponding 3D representations



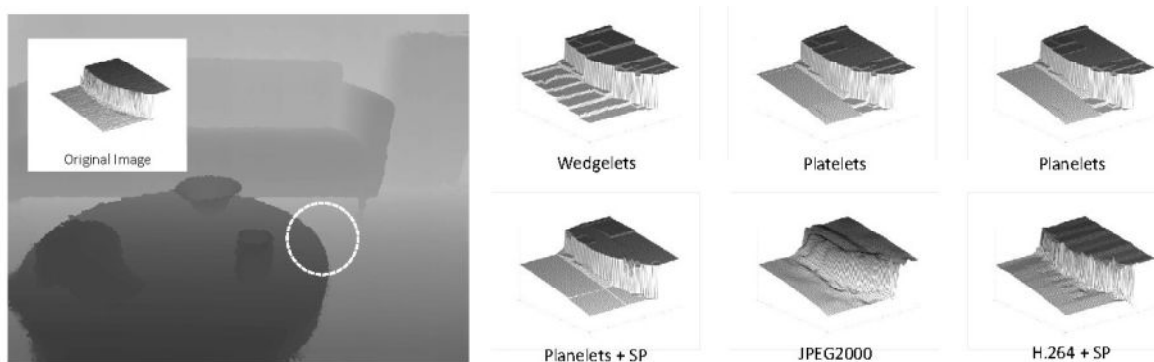
(شکل-۱۱): نمای سه‌بعدی صحنه حاصل از بازسازی یک تصویر عمق مصنوعی با ساختار هندسی صفحه‌گون در نرخ بیت 0.02 bpp  
(Figure-11): 3D scene generated by depth image of a synthetic indoor scene with planar geometry decoded at the bit-rate of 0.02 bpp



(شکل-۱۲): نمای سه‌بعدی صحنه حاصل از بازسازی یک تصویر عمق واقعی با ساختار هندسی صفحه‌گون در نرخ بیت 0.02 bpp  
(Figure-12): 3D scene generated by depth image of a real-world indoor scene with planar geometry decoded at the bit-rate of 0.02 bpp



(شکل-۱۳): یک لبه تیز در تصویر عمق فاقد نوفه که توسط کدگذارهای مختلف در نرخ بیت 0.03 bpp بازسازی شده است  
(Figure-13): Surface plots of a sharp edge in a synthetic depth image decoded by different compression schemes at the bit-rate of 0.03 bpp



(شکل-۱۴): یک لبه تیز در تصویر عمق کینکت که توسط کدگذارهای مختلف در نرخ بیت 0.03 bpp بازسازی شده است  
(Figure-14): Surface plots of a sharp edge in a Kinect-like depth image decoded by different compression schemes at the bit-rate of 0.03 bpp

کیفیت حفظ تصویر در کاربردهای فشرده‌سازی است که به‌شکل زیر بر روی مقادیر عمق تعریف شده و حین ارزیابی تفاوتی بین اجزای تصویر قایل نمی‌شود:

$$PSNR_Z = 10 \cdot \log_{10} \left( \frac{P_Z^2}{MSE_Z} \right) \quad (21)$$

## ۸-۲- ارزیابی کمی کیفیت تصویر

به‌منظور ارزیابی و مقایسه عددی کیفیت تصاویر بازسازی‌شده توسط روش‌های فشرده‌سازی مختلف، در این مقاله از معیار کیفیت PSNR سنتی و معیار کیفیت SSIM استفاده شد. معیار کیفیت PSNR سنتی رایج‌ترین معیار برای ارزیابی

(جدول-۱): میانگین کیفیت تصاویر بازسازی شده در نرخ بیت

۰/۰۳ bpp برای کدگذارهای مختلف تصویر

(Table-1): Average quality of reconstructed depth images at the bit-rate of 0.03 bpp for different compression algorithms

SSIM واقعی (dB)	SSIM مصنوعی (dB)	PSNR واقعی (dB)	PSNR مصنوعی (dB)	کدگذار تصویر
0.1330	0.1461	42.32	51.68	Wedgelets
0.4425	0.7625	43.54	57.77	Platelets
0.4436	0.7754	43.80	58.60	Planelets
<b>0.4839</b>	<b>0.8385</b>	<b>45.00</b>	<b>63.50</b>	Planelets + SP
0.3637	0.5523	42.41	54.00	JPEG2000
0.3690	0.7105	43.44	58.94	H.264 + SP

### ۳-۸- ارزیابی زمان مصرفی

کدگذاری تصاویر عمق به کمک موجک‌های هندسی نیازمند ساخت درخت چهارگانه، تقریب هر بلاک به کمک بهترین عنصر دیکشنری، هرس درخت برای گام‌های چندی‌سازی مختلف، و کدگذاری بهترین درخت هرس شده در یک جریان بیتی است. به علاوه، برای محاسبه تبدیلات مختلف در گروه موجک‌های هندسی وفقی نمی‌توان از کانولوشن یا تبدیل فوریه استفاده کرد. این عوامل باعث شده‌اند کدگذاری تصاویر به کمک موجک‌های هندسی وفقی نسبت به کدگذاری تصاویر به کمک موجک سنتی و DCT به زمان بیشتری نیاز داشته باشد.

در این آزمایش، میانگین زمان مصرفی کدگذارهای مبتنی بر موجک‌های هندسی و کدگذارهای مدرن تصویر برای فشرده‌سازی تصاویر عمق کینکت با اندازه  $480 \times 640$  ارزیابی شد. این آزمایش بر روی یک رایانه مجهز به پردازنده Intel Core i7 3.5 GHz و حافظه اصلی 8 GB انجام شد. نتایج ارزیابی زمان مصرفی برای کدگذارهای مختلف در جدول (۲) خلاصه شده است. همان‌طور که در این جدول مشاهده می‌شود کدگذارهای مدرن تصویر که بر تبدیل موجک سنتی و DCT مبتنی هستند، به دلیل استفاده از کانولوشن برای محاسبه ضرایب تبدیل، سرعت بسیار خوبی داشته و به کمتر از یک ثانیه زمان برای فشرده‌سازی هر تصویر نیاز دارند. در مقابل، کدگذارهای مبتنی بر موجک‌های هندسی به دلیل عدم استفاده از کانولوشن برای محاسبه ضرایب، به بیش از چند ثانیه زمان برای فشرده‌سازی هر تصویر نیاز دارند.

از آنجایی که محاسبه تبدیل صفحک بر پایه تبدیل گوک انجام می‌شود، زمان مصرفی کدگذار مبتنی بر صفحک

که در آن  $P_z$  نشان‌دهنده بیشینه مقدار عمق ممکن برای تصاویر است (در مجموعه تصاویر این مقاله بیشینه مقدار عمق ۶۵۵۳۵ است). در مقابل، معیار کیفیت  $SSIM^1$  به حفظ کیفیت ساختارهای تصویر از جمله لبه‌ها اهمیت بیشتری می‌دهد. به جای مقایسه پیکسل به پیکسل تصویر بازسازی شده با تصویر اصلی، معیار  $SSIM$  در پنجره‌های محلی میزان انطباق میانگین، واریانس، و کوواریانس مقادیر پیکسل‌ها را به یکدیگر می‌سنجد [43].

میانگین کیفیت تصاویر بازسازی شده برای هر دو مجموعه تصویر عمق مصنوعی و واقعی در نرخ بیت متوسط ۰/۰۳ bpp برای الگوریتم‌های مختلف در جدول (۱) خلاصه شده است. همان‌طور که در این جدول مشاهده می‌شود، از نظر معیار  $PSNR$  سنتی، استفاده از مدل غیرخطی برای بازسازی نواحی صفحه‌گون در Planelets باعث ۶/۹ dB افزایش کیفیت برای تصاویر عمق مصنوعی، و ۱/۴۵ dB افزایش کیفیت تصاویر عمق واقعی نسبت به Wedgelets شده است. استفاده از پیش‌بینی مکانی در Planelets+SP باعث شده است تا بهبود کیفیت حاصل از روش پیشنهادی نسبت به Wedgelets برای تصاویر عمق مصنوعی به ۱۱/۸ dB و برای تصاویر عمق واقعی به ۲/۷ dB افزایش یابد. در مقایسه با تبدیل شیبک، استفاده از مدل غیرخطی متناظر با تغییرات عمق نواحی صفحه‌گون در تبدیل صفحک باعث شده است تا میانگین کیفیت بر روی تصاویر عمق مصنوعی از ۵۷/۷۷ dB به ۵۸/۶۰ dB و بر روی تصاویر عمق واقعی از ۴۳/۵۴ dB به ۴۳/۸۰ dB افزایش یابد. در مقایسه با روش‌های فشرده‌سازی مدرن، استفاده از روش پیشنهادی این مقاله همراه با روش‌های پیش‌بینی مکانی منجر به ۲/۵۹ dB افزایش کیفیت نسبت به JPEG2000 و ۱/۵۶ dB افزایش کیفیت نسبت به H.264/AVC برای تصاویر عمق واقعی شده است.

از نظر حفظ کیفیت لبه‌ها نیز نتایج معیار  $SSIM$  نشان می‌دهد که استفاده از مدل‌های غیرثابت در موجک‌های هندسی شیبک و صفحک باعث شده است تا فشرده‌سازی تصاویر عمق به کمک آن‌ها نسبت به روش‌های عمومی JPEG2000 و H264 عملکرد بسیار بهتری را از نظر حفظ ساختارهای تصویر داشته باشد. همچنین از نظر معیار  $SSIM$  نیز روش پیشنهادی این مقاله، یعنی تبدیل صفحک درجه‌دو در میان روش‌های مقایسه‌شده بهترین عملکرد را در حفظ جزئیات تصویر داشته است.

<sup>1</sup> Structural Similarity Index Measure (SSIM)

بهره‌گیری از پیش‌بینی بین‌بلاکی به زمان مصرفی کدگذار مبتنی بر صفحک در فشرده‌سازی جریان‌های عمق، و تسریع عملکرد آن به کمک محاسبات موازی مهم‌ترین مسیرهای پیشنهادی ما برای ادامه این پژوهش هستند.

## پیوست ۱- نحوه کمینه‌سازی رابطه تقریب صفحک

در بخش ۱-۵- نحوه محاسبه تبدیل گوک تکراری برای هر بلاک از تصویر تشریح شد، که به محاسبه پارامترهای لبه برای هر بلاک به صورت سه‌تایی مرتب  $(v_1, v_2, d)$  منجر شد. همچنین در بخش ۲-۵- اشاره شد که لبه پیدا شده باید با دو تابع گویا با رابطه (۱۰) تزئین شود تا بلاک ورودی توسط تبدیل صفک درجه دو تقریب زده شود. از آنجایی که پارامترهای لبه یعنی  $(v_1, v_2, d)$  در بخش ۱-۵- محاسبه شده‌اند، این پارامترها در بخش ۲-۵- به عنوان ثابت در نظر گرفته می‌شوند؛ لذا به منظور تزئین لبه با دو تابع گویا تنها کافی است پارامترهای دو تابع گویا به صورت چندتایی مرتب  $(\alpha_1, \beta_1, \gamma_1)$  و چندتایی مرتب  $(\alpha_2, \beta_2, \gamma_2)$  بر اساس محتوای بلاک تخمین زده شوند. بدین منظور باید رابطه (۱۷) برای پیکسل‌های بلاک جاری و لبه  $(v_1, v_2, d)$  کمینه‌سازی شود. از آنجایی که مکان لبه ثابت است؛ لذا در این فرآیند کمینه‌سازی دو ناحیه حاصل از لبه در بلاک جاری که با نام  $R$  و  $\bar{R}$  به آن‌ها اشاره شد، در طول کمینه‌سازی تغییر نخواهند کرد. به همین دلیل، برای کمینه‌سازی رابطه (۱۷) کافی است دو جمله آن جدا شده و کمینه‌سازی شوند. کمینه‌سازی جمله نخست رابطه (۱۷) معادل با کمینه‌سازی معادله زیر برای پیکسل‌های ناحیه  $R$  خواهد بود:

$$\sum_{(x_i, y_i) \in R} \left( (\alpha_1 x + \beta_1 y + \gamma_1) - \frac{1}{I(x_i, y_i)} \right)^2 \quad (22)$$

به‌طور مشابه، کمینه‌سازی جمله دوم رابطه (۱۶) معادل با کمینه‌سازی معادله زیر برای پیکسل‌های ناحیه  $\bar{R}$  خواهد بود:

$$\sum_{(x_i, y_i) \in \bar{R}} \left( (\alpha_2 x + \beta_2 y + \gamma_2) - \frac{1}{I(x_i, y_i)} \right)^2 \quad (23)$$

به منظور کمینه‌سازی رابطه (۲۲) کافی است از آن نسبت به هر یک از پارامترهای  $\alpha_1$  و  $\beta_1$  و  $\gamma_1$  مشتق گرفته و برابر با صفر قرار دهیم، تا به دستگاه معادلات زیر برای محاسبه این پارامترها برسیم:

بدون استفاده از پیش‌بینی بین‌بلاکی به زمان مصرفی کدگذار مبتنی بر گوک نزدیک بوده و تنها یک ثانیه بیشتر است. هنگامی که از پیش‌بینی بین‌بلاکی نیز در کدگذار مبتنی بر صفحک استفاده می‌شود، به دلیل اینکه پیش‌بینی بین‌بلاکی باید حین فرآیند هرس انجام شود، زمان مصرفی الگوریتم پیشنهادی حدود شش ثانیه افزایش می‌یابد. طرح الگوریتم تبدیل گوک تکراری در این مقاله به محاسبه سریع‌تر موجک‌های هندسی نسبت به الگوریتم جستجوی جامع منجر شد، و امکان استفاده از موجک‌های هندسی را در کاربردهای آفلاین با زمان قابل تحمل فراهم آورد. با این وجود، هنوز زمان مصرفی این الگوریتم برای ارسال جریانی از تصاویر عمق به صورت بلادرنگ قابل تحمل نیست.

(جدول ۲): میانگین زمان مصرفی کدگذارهای مختلف برای

فشرده‌سازی یک تصویر عمق با اندازه  $640 \times 480$

(Table-2): Average running time (sec.) of different compression algorithms for compressing a depth image of size  $640 \times 480$

الگوریتم فشرده‌سازی	زمان مصرفی بر حسب ثانیه
Wedgelets	5.80
Platelets	6.19
Planelets	6.72
Planelets + SP	12.10
JPEG2000	0.08
H.264 + SP	0.25

## ۹- جمع‌بندی و نتیجه

در این مقاله تبدیل صفحک به عنوان عضو جدیدی از خانواده موجک‌های هندسی وقتی برای بازنمایی کارآمد تصاویر عمق گرفته شده از محیط‌های داخلی معرفی و سپس الگوریتم کدگذار مبتنی بر صفحک برای فشرده‌سازی تصاویر عمق به کمک تبدیل صفحک مطرح و عملکرد آن ارزیابی شد.

کدگذار عمق مبتنی بر صفحک به دلیل بازنمایی صریح لبه‌های تیز در تصویر و استفاده از مدل صفحه‌گون برای بازنمایی سطوح نسبت به موجک‌های هندسی گوک و شیبک به کیفیت بالاتری در نرخ بیت مشابه دست می‌یابد؛ به علاوه، کدگذار عمق مبتنی بر صفحک نسبت به کدگذارهای تصاویر عمومی در حفظ هندسه صحنه توانا تر بوده و در مقابل نوفه نیز مقاوم‌تر است. با این وجود، به دلیل عدم استفاده از کانولوشن برای محاسبه ضرایب در کدگذارهای مبتنی بر موجک‌های هندسی، زمان مصرفی این کدگذارها هنوز نسبت به کدگذارهای مدرن JPEG2000 و H.264 بیشتر است.

Mutto, P. Zanuttigh, and G. Cortelazzo, Eds., cd: Springer US, 2012, pp. 33-47.

- [8] J. Fu, D. Miao, W. Yu, S. Wang, Y. Lu, and S. Li, "Kinect-like depth data compression," *IEEE Transactions on Multimedia*, vol. 15, pp. 1340 - 1352, October 2013.
- [9] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Müller, P. de With, et al., "The effects of multiview depth video compression on multiview rendering," *Signal Processing: Image Communication*, vol. 24, pp. 73-88, January 2009.
- [10] K. Muller, P. Merkle, and T. Wiegand, "3-D Video Representation Using Depth Maps," *Proceedings of the IEEE*, vol. 99, pp. 643-656, April 2010.
- [11] J. Ruiz-Hidalgo, J. Morros, P. Aflaki, F. Calderero, and F. Marqués, "Multiview depth coding based on combined color/depth segmentation," *Journal of Visual Communication and Image Representation*, vol. 23, pp. 42-52, January 2012.
- [12] I. Daribo, H. Saito, R. Furukawa, S. Hiura, and N. Asada, "Effects of Wavelet-Based Depth Video Compression," in *3D-TV System with Depth-Image-Based Rendering*, C. Zhu, Y. Zhao, L. Yu, and M. Tanimoto, Eds., cd: Springer New York, 2013, pp. 277-298.
- [13] M. Maitre, and M. Do, "Depth and depth-color coding using shape-adaptive wavelets," *Journal of Visual Communication and Image Representation*, vol. 21, pp. 513-522, July-August 2010.
- [14] D. Donoho, "Wedgelets: nearly minimax estimation of edges," *Annals of Statistics*, vol. 27, pp. 859-897, April 1999.
- [15] A. Lisowska, *Geometrical Multiresolution Adaptive Transforms - Theory and Applications*, 1st ed.: Springer International Publishing, 2014.
- [16] J. Romberg, M. Wakin, and R. Baraniuk, "Multiscale wedgelet image analysis: fast decompositions and modeling," in *International Conference on Image Processing*, Rochester, New York, 2002, pp. 585-588.
- [17] H. Bagherzadeh, A. Harati, Z. Amiri, and R. KamyabiGol, "Video Denoising Using block Shearlet Transform," *Signal and Data Processing*, vol. 15, pp. 17-30, 2018.
- [18] R. Willett and R. Nowak, "Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging," *IEEE Transactions on Medical Imaging*, vol. 22, pp. 332-350, March 2003.

[۱۷] ح. باقرزاده، ا. هراتی، ز. امیری، و ر. کامیابی گل، «رفع نوفه ویدئو توسط تبدیل قیچک قطعه‌ای»، پردازش علائم و داده‌ها، دوره ۱۵، شماره ۲، ۱۳۹۷.

$$\begin{pmatrix} \sum_i x_i^2 & \sum_i x_i y_i & \sum_i x_i \\ \sum_i x_i y_i & \sum_i y_i^2 & \sum_i y_i \\ \sum_i x_i & \sum_i y_i & \sum_i 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \beta_1 \\ \gamma_1 \end{pmatrix} = \begin{pmatrix} \sum_i x_i \frac{1}{I(x_i, y_i)} \\ \sum_i y_i \frac{1}{I(x_i, y_i)} \\ \sum_i \frac{1}{I(x_i, y_i)} \end{pmatrix} \quad (۲۳)$$

که در آن هر  $(x_i, y_i)$  یک نقطه از ناحیه  $R$  است. مشابه این دستگاه معادلات برای ناحیه  $\bar{R}$  و پارامترهای  $\alpha_1$  و  $\beta_1$  و  $\gamma_1$  نیز قابل نوشتن است. دستگاه معادلات خطی بالا می‌تواند با محاسبه معکوس ماتریس حل شود. از آنجایی که عملیات معکوس ماتریس در دستگاه معادلات فوق همواره باید بر روی یک ماتریس  $3 \times 3$  انجام شود، پیچیدگی محاسباتی حل دستگاه معادلات بالا قابل تحمل خواهد بود، و ارتباطی به تعداد پیکسل‌های بلاک نخواهد داشت.

## 10- References

## ۱۰- مراجع

- [1] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," *International Journal of Robotics Research*, vol. 31, pp. 647-663, April 2012.
- [2] K. Khoshelham, and S. Elberink, "Accuracy and Resolution of Kinect Depth Data for Indoor Mapping Applications," *Sensors*, vol. 12, pp. 1437-1454, February 2012.
- [3] Y. Wang, F. Zhong, Q. Peng, and X. Qin, "Depth map enhancement based on color and depth consistency," *The Visual Computer*, vol. 30, pp. 1157-1168, October 2014.
- [4] T. Whelan, H. Johannsson, M. Kaess, J. Leonard, and J. McDonald, "Robust real-time visual odometry for dense RGB-D mapping," in *IEEE International Conference on Robotics and Automation (ICRA)*, Karlsruhe, Germany, 2013, pp. 5724-5731.
- [۵] م. پاک نژاد، م. رضائیان، «مدل سازی صفحه‌ای محیط‌های داخلی با استفاده از تصاویر RGB-D»، پردازش علائم و داده‌ها، دوره ۱۴، شماره ۳، ۱۳۹۶.
- [5] M. Paknezhad, and M. Rezaeian, "Indoor Planar Modeling Using RGB-D Images," *Signal and Data Processing*, vol. 14, pp. 143-160, 2017.
- [6] J. Smisek, M. Jancosek, and T. Pajdla, "3D with Kinect," in *IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, Barcelona, Spain, 2011, pp. 1154-1160.
- [7] C. Mutto, P. Zanuttigh, and G. Cortelazzo, "Microsoft Kinect™ Range Camera," in *Time-of-Flight Cameras and Microsoft Kinect™*, C.

- Prediction for 3-D Video Systems," *IEEE Transactions on Multimedia*, vol. 14, pp. 121-128, February 2012.
- [30] F. Pecc, J. Kautz, and T. Weyrich, "Three Depth-Camera Technologies Compared," in *1st BEAMING Workshop*, Barcelona, Spain, 2011, pp. 1-4.
- [31] S. Mehrotra, Z. Zhengyou, C. Qin, Z. Cha, and P. Chou, "Low-complexity, near-lossless coding of depth maps from kinect-like depth cameras," in *IEEE 13th International Workshop on Multimedia Signal Processing (MMSP)*, Saint-Malo, France, 2011, pp. 1-6.
- [32] D. Sandberg, P. Forssen, and J. Ogniewski, "Model-Based Video Coding Using Colour and Depth Cameras," in *International Conference on Digital Image Computing Techniques and Applications (DICTA)*, Noosa, QLD, Australia, 2011, pp. 158-163.
- [33] D. Donoho and X. Huo, "Beamlets and Multiscale Image Analysis," in *Multiscale and Multi-resolution Methods*, vol. 20, T. Barth, T. Chan, and R. Haimes, Eds., ed: Springer Berlin Heidelberg, 2002, pp. 149-196.
- [34] A. Lisowska, "Second Order Wedgelets in Image Coding," in *The International Conference on Computer as a Tool*, Warsaw, Poland, 2007, pp. 237-244.
- [35] A. Lisowska, "Moments-Based Fast Wedgelet Transform," *Journal of Mathematical Imaging and Vision*, vol. 39, pp. 180-192, February 2011.
- [36] A. Lisowska, "Smoothlet Transform: Theory and Applications," in *Advances in Imaging and Electron Physics*, vol. 178, W. H. Peter, Ed., ed: Elsevier, 2013, pp. 97-145.
- [37] F. Friedrich, L. Demaret, H. Führ, and K. Wicker, "Efficient Moment Computation over Polygonal Domains with an Application to Rapid Wedgelet Approximation," *SIAM Journal on Scientific Computing*, vol. 29, pp. 842-863, April 2007.
- [38] V. Kiani, A. Harati, and A. Vahedian, "Iterative Wedgelet Transform: An efficient algorithm for computing wedgelet representation and approximation of images," *Journal of Visual Communication and Image Representation*, vol. 34, pp. 65-77, January 2016.
- [39] V. Kiani, A. Harati, and A. Vahedian, "A relaxation approach to computation of second-order wedgelet transform with application to image compression," *Signal Processing: Image Communication*, vol. 41, pp. 115-127, February 2016.
- [40] R. Shukla, P. Dragotti, M. Do, and M. Vetterli, "Rate-distortion optimized tree-structured compression algorithms for piecewise polynomial images," *IEEE Transactions on Image Processing*, vol. 14, pp. 343-359, March 2005.
- [19] M. Dou, L. Guan, J. Frahm, and H. Fuchs, "Exploring High-Level Plane Primitives for Indoor 3D Reconstruction with a Hand-held RGB-D Camera," in *Computer Vision - ACCV 2012 Workshops*, vol. 7729, J.-I. Park and J. Kim, Eds., ed: Springer Berlin Heidelberg, 2013, pp. 94-108.
- [20] R. Kaushik and J. Xiao, "Accelerated patch-based planar clustering of noisy range images in indoor environments for robot mapping," *Robotics and Autonomous Systems*, vol. 60, pp. 584-598, April 2012.
- [21] M. Brown, D. Burschka, and G. Hager, "Advances in computational stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 993-1008, August 2003.
- [22] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7-42, April-June 2002.
- [23] N. Thakoor, J. Sungying, and G. Jean, "Real-time Planar Surface Segmentation in Disparity Space," in *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, USA, 2007, pp. 1-8.
- [24] D. Sebai, F. Chaieb, K. Mammou, and F. Ghorbel, "Piece-wise linear function estimation for platelet-based depth maps coding using edge detection," in *Proceedings of SPIE 8290, Three-Dimensional Image Processing (3DIP) and Applications II*, Burlingame, California, United States, 2012, p. 82901C.
- [25] Z. Yu, W. Han, and Z. Jiying, "Depth map compression based on platelet coding and quadratic curve fitting," in *IEEE 27th Canadian Conference on Electrical and Computer Engineering (CCECE)*, Halifax, Nova Scotia, Canada, 2014, pp. 1-4.
- [26] I. Daribo, C. Tillier, and B. Pesquet-Popescu, "Adaptive wavelet coding of the depth map for stereoscopic view synthesis," in *IEEE 10th Workshop on Multimedia Signal Processing*, Cairns, Queensland, Australia, 2008, pp. 413-417.
- [27] O. Kwan-Jung, Y. Schoon, A. Vetro, and H. Yo-Sung, "Depth Reconstruction Filter and Down/Up Sampling for Depth Coding in 3-D Video," *IEEE Signal Processing Letters*, vol. 16, pp. 747-750, September 2009.
- [28] O. Kwan-Jung, A. Vetro, and H. Yo-Sung, "Depth Coding Using a Boundary Reconstruction Filter for 3-D Video Systems," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, pp. 350-359, March 2011.
- [29] K. Min-Koo and H. Yo-Sung, "Depth Video Coding Using Adaptive Geometry Based Intra

پژوهشی مورد علاقه ایشان عبارتند از: پردازش تصویر و ویدئو،  
بینایی ماشین، و سیستم‌های هوشمند ترافیک.  
نشانی رایانامه ایشان عبارت است از:

vahedian@um.ac.ir

- [41] A. Handa, T. Whelan, J. McDonald, and A. Davison, "A Benchmark for RGB-D Visual Odometry, 3D Reconstruction and SLAM," in *IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, 2014, pp. 1524-1531.
- [42] K. Lai, L. Bo, and D. Fox, "Unsupervised feature learning for 3D scene labeling," in *IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, 2014, pp. 3050-3057.
- [43] W. Zhou, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600-612, 2004.

#### وحید کیانی مدرک کارشناسی ارشد و



دکترای خود را به ترتیب در سال‌های ۱۳۹۰ و ۱۳۹۵ از دانشگاه فردوسی مشهد در رشته مهندسی کامپیوتر دریافت کرده است. در حال حاضر، ایشان عضو هیئت علمی و

استادیار دانشکده مهندسی دانشگاه بجنورد است. زمینه‌های پژوهشی مورد علاقه ایشان عبارتند از: موجک‌های سنتی و هندسی، تبدیلات چنددقتی، فشرده‌سازی تصویر و ویدئو، و یادگیری ماشین. این مقاله حاصل پژوهش وی در زمان کار روی پایان‌نامه دکترای در دانشگاه فردوسی مشهد است. نشانی رایانامه ایشان عبارت است از:

v.kiani@ub.ac.ir

#### احد هراتی مدرک کارشناسی ارشد خود را



در رشته هوش مصنوعی و رباتیک در سال ۱۳۸۱ از دانشگاه تهران دریافت کرده است. همچنین درجه دکترای خود را از انستیتو تکنولوژی فدرال زوریخ (ETHZ) در سال

۱۳۸۶ اخذ کرده است. در حال حاضر، ایشان عضو هیئت علمی و استادیار گروه مهندسی کامپیوتر دانشگاه فردوسی مشهد است. زمینه‌های پژوهشی مورد علاقه ایشان عبارتند از: ادراک ربات، بینایی سه‌بعدی، و مدل‌سازی احتمالاتی. نشانی رایانامه ایشان عبارت است از:

a.harati@um.ac.ir

#### عابدین واحدیان درجه دکترای خود را از



دانشگاه نیو ساوت ولز (UNSW) استرالیا در سال ۱۳۷۷ اخذ کرده است. در حال حاضر، ایشان عضو هیئت علمی و دانشیار گروه

مهندسی کامپیوتر دانشگاه فردوسی مشهد است. زمینه‌های